

ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ

ΔΥΤΙΚΗΣ ΕΛΛΑΔΑΣ

ΣΧΟΛΗ ΔΙΟΙΚΗΣΗΣ & ΟΙΚΟΝΟΜΙΑΣ

ΤΜΗΜΑ ΔΙΟΙΚΗΣΗΣ ΕΠΙΧΕΙΡΗΣΕΩΝ



ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

**«ΕΡΓΑΛΕΙΑ ΑΝΑΛΥΣΗΣ ΚΑΙ ΟΠΤΙΚΟΠΟΙΗΣΗΣ
ΔΕΔΟΜΕΝΩΝ ΣΕ ΣΥΣΤΗΜΑΤΑ
ΕΠΙΧΕΙΡΗΜΑΤΙΚΗΣ ΕΥΦΥΪΑΣ»**

Διένης Παναγιώτης

Μπολοβίνος Αλέξανδρος

ΕΙΣΗΓΗΤΗΣ

Δελτούζος Κωνσταντίνος Ακαδημαϊκός υπότροφος

ΜΕΣΟΛΟΓΓΙ 2019

Η έγκριση της πτυχιακής εργασίας από το Τμήμα Διοίκησης Επιχειρήσεων / Μεσολογγίου του ΤΕΙ Δυτικής Ελλάδας δεν υποδηλώνει απαραίτητως και αποδοχή των απόψεων του συγγραφέα εκ μέρους του Τμήματος.

Πρόλογος-Ευχαριστίες

Η επιστήμη / τεχνολογία της ανάλυσης δεδομένων, δημιουργήθηκε τα τελευταία χρόνια με την έλευση της πληροφορικής και του προβλήματος διαχείρισης, αποθήκευσης αλλά και άμεσης πρόσβασης σε τεράστιες βάσεις δεδομένων. Μάλιστα έχουν δημιουργηθεί - από τους επιστήμονες που σχετίζονται με το θέμα- διάφορες προσεγγίσεις, οι οποίες έχουν δημιουργήσει διαφορετικές τεχνικές, σε διαφορετικούς επιχειρηματικούς τομείς, επιστημονικούς αλλά και κοινωνικούς. Βέβαια, στη σημερινή επιχειρηματική δραστηριότητα, ο όρος «ανάλυση δεδομένων» διαδραματίζει βασικό ρόλο στη λήψη επιχειρηματικών αποφάσεων με επενέργεια στην αποτελεσματική λειτουργία της εκάστοτε επιχείρησης.

Το θέμα της παρούσης πτυχιακής εργασίας είναι τα: «Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας» και εγγράφηκε από τους φοιτητές: Διένη Παναγιώτη και Μπολοβίνο Αλέξανδρο, στα πλαίσια του προπτυχιακού προγράμματος του τμήματος Διοίκησης Επιχειρήσεων της σχολής Διοίκησης & Οικονομίας του Ανώτατου Τεχνολογικού Εκπαιδευτικού Ιδρύματος Δυτικής Ελλάδας.

Σε αυτό το σημείο θα θέλαμε να ευχαριστήσουμε από καρδιάς, τον καθηγητή και εισηγητή του θέματος, κύριο Δελτούζο Κωνσταντίνο, για την αμέριστη επιστημονική και ηθική συμπαράσταση και βοήθεια στην ολοκλήρωση της παρούσας πτυχιακής εργασίας!

Όλους τους καθηγητές του τμήματός μας, για τις γνώσεις που μας προσέφεραν στα χρόνια των σπουδών μας.

Τέλος θα θέλαμε να ευχαριστήσουμε και να αφιερώσουμε την πτυχιακή εργασία στους γονείς μας για όσα έκαναν και κάνουν για μας.

Διένης Παναγιώτης

Μπολοβίνος Αλέξανδρος

Περίληψη

Στην παρούσα πτυχιακή εργασία αρχικά αναφερόμαστε στην διαδικασία της «Ανάλυσης δεδομένων». Προτάσσονται γενικά στοιχεία όπως, η διαδικασία ανάλυσης δεδομένων, τα ποσοτικά μηνύματα, οι τεχνικές ανάλυσης ποσοτικών δεδομένων, τα προβλήματα αποτελεσματικής ανάλυσης αλλά και τα ελεύθερα λογισμικά για την ανάλυση δεδομένων. Πραγματοποιείται ανάλυση της διεργασίας «Οπτικοποίηση δεδομένων» και παράλληλα παρατίθενται διάφοροι ορισμοί / έννοιες, τα χαρακτηριστικά των αποτελεσματικών γραφικών απεικονίσεων, τα ποσοτικά μηνύματα, η οπτική αντίληψη και απεικόνιση δεδομένων, η ιστορική αναδρομή της απεικόνισης δεδομένων και διάφορα παραδείγματα διαγραμμάτων χρήσης οπτικοποίησης δεδομένων αλλά και η αρχιτεκτονική παρουσίασης δεδομένων. Έπειτα αποσαφηνίζεται η έννοια της «Επιχειρηματικής ευφυΐας», αναλύονται οι λειτουργίες και τεχνολογίες της επιχειρηματικής ευφυΐας, τα χαρακτηριστικά της, τα δεδομένα επιχειρηματικών λειτουργιών, τα μεταδεδομένα και οι εφαρμογές επιχειρηματικής ευφυΐας. Εν συνεχεία στοιχειοθετείται η «Επιχειρηματική ευφυΐα σε πραγματικό χρόνο», δίνοντας τον σκοπό της, τα τεχνολογικά στοιχεία επιχειρηματικής ευφυΐας, την επιχειρηματική νοημοσύνη σε πραγματικό χρόνο, την λήψη αποφάσεων επιχειρηματικής ευφυΐας, τις αρχιτεκτονικές επιχειρηματικής ευφυΐας και τις τεχνολογίες που υποστηρίζουν την ανάλυση σε πραγματικό χρόνο. Τέλος παρουσιάζονται τα «Εργαλεία οπτικοποίησης δεδομένων», καταγράφονται γενικά στοιχεία για τα εργαλεία οπτικοποίησης δεδομένων, δηλαδή για τα: Tableau, Qlikview, FusionCharts, Highcharts, Datawrapper, Plotly, Sisense, Tableau Public, Microsoft Power BI.

Abstract

In this thesis we initially refer to the «Data Analysis» process. Generally, data such as data analysis, quantitative messages, quantitative data analysis techniques, effective analysis problems, and free software for data analysis are proposed. An analysis of the «Data Visualization» process is carried out, and various definitions / concepts, the characteristics of effective graphical representations, quantitative messages, visual perception and data display, historical data visualization and various examples of data visualization graphs, as well as the data presentation architecture. Then the concept of Business Intelligence is clarified, the functions and technologies of business intelligence, its features, business functions data, metadata and business intelligence applications are analyzed. Real Business Intelligence is then set up, providing business intelligence, business intelligence, real-time business intelligence, business intelligence, business intelligence architectures, and technologies that support real-time analysis. Finally, the data visualization tools are presented, data for data visualization tools, ie: Tableau, Qlikview, FusionCharts, Highcharts, Datawrapper, Plotly, Sisense, Tableau Public, and Microsoft Power BI.

Πινάκας περιεχομένων

Πρόλογος-Ευχαριστίες	iii
Περίληψη.....	iv
Abstract	v
Πινάκας περιεχομένων	vi
Κατάλογος Σχημάτων - Πινάκων.....	x
Συντομογραφίες - Απόδοση όρων.....	xi
1 Κεφάλαιο: «Εισαγωγή»	1
2 Κεφάλαιο: «Ανάλυση δεδομένων».....	3
2.1 Γενικά στοιχεία	3
2.2 Διαδικασία ανάλυσης δεδομένων	4
2.2.1 Απαιτήσεις δεδομένων	5
2.2.2 Συλλογή δεδομένων	6
2.2.3 Επεξεργασία δεδομένων.....	6
2.2.4 Καθαρισμός δεδομένων	6
2.2.5 Διερευνητική ανάλυση δεδομένων.....	7
2.2.6 Μοντελοποίηση και αλγόριθμοι.....	7
2.3 Ποσοτικά μηνύματα	8
2.4 Τεχνικές ανάλυσης ποσοτικών δεδομένων	9
2.5 Προβλήματα αποτελεσματικής ανάλυσης	12
2.5.1 Γεγονός που προκαλεί σύγχυση.....	12
2.5.2 Γνωσιακές προκαταλήψεις.....	12
2.5.3 Αναξιοπιστία	13
2.6 Ελεύθερα λογισμικά για την ανάλυση δεδομένων.....	14
3 Κεφάλαιο: «Οπτικοποίηση δεδομένων».....	20

3.1	Γενικά στοιχεία	20
3.2	Ορισμοί / έννοιες.....	20
3.3	Χαρακτηριστικά των αποτελεσματικών γραφικών απεικονίσεων.....	21
3.4	Ποσοτικά μηνύματα	22
3.5	Οπτική αντίληψη και απεικόνιση δεδομένων	24
3.6	Ιστορική αναδρομή της απεικόνισης δεδομένων	25
3.7	Έννοιες – ορισμοί.....	26
3.8	Παραδείγματα διαγραμμάτων χρήσης οπτικοποίησης δεδομένων	27
3.8.1	Ραβδόγραμμα	27
3.8.2	Ιστόγραμμα.....	28
3.8.3	Διάγραμμα διασποράς.....	30
3.8.4	Διάγραμμα έργου	31
3.8.5	Streamgraph	32
3.8.6	Treemapping.....	33
3.8.7	Heatmap	34
3.8.8	Διαγράμματα Spider.....	35
3.9	Αρχιτεκτονική παρουσίασης δεδομένων	36
4	Κεφάλαιο: «Επιχειρηματική ευφυΐα»	39
4.1	Λειτουργίες και τεχνολογίες επιχειρηματικής ευφυΐας.....	39
4.2	Χαρακτηριστικά Επιχειρηματικής ευφυΐας.....	42
4.2.1	Σύγκριση με την ανταγωνιστική νοημοσύνη	43
4.2.2	Σύγκριση με τις επιχειρηματικές αναλύσεις	43
4.3	Στοιχεία της Επιχειρηματικής ευφυΐας	43
4.4	Δεδομένα επιχειρηματικών λειτουργιών.....	44
4.4.1	Μη δομημένα δεδομένα έναντι ημιδομημένων δεδομένων	45
4.4.2	Περιορισμοί ημιδομημένων και αδόμητων δεδομένων	46

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας	
4.5	Μεταδεδομένα..... 46
4.6	Εφαρμογές Επιχειρηματικής ευφυΐας 47
4.6.1	Μετρήσεις απόδοσης..... 47
4.6.2	Συγκριτική αξιολόγηση 49
4.6.3	Διαχείριση γνώσης 50
5	Κεφάλαιο: «Επιχειρηματική ευφυΐα σε πραγματικό χρόνο» 52
5.1	Γενικά στοιχεία 52
5.2	Σκοπός Επιχειρηματικής Ευφυΐας..... 52
5.3	Διάφορα στοιχεία Επιχειρηματικής Ευφυΐας 53
5.4	Τεχνολογικά στοιχεία Επιχειρηματικής Ευφυΐας..... 53
5.5	Επιχειρηματική νοημοσύνη σε πραγματικό χρόνο 55
5.6	Λήψη αποφάσεων Επιχειρηματικής ευφυΐας 56
5.7	Καθυστέρηση Επιχειρηματικής ευφυΐας..... 56
5.8	Αρχιτεκτονικές Επιχειρηματικής ευφυΐας..... 57
5.8.1	Βάση συμβάντων..... 57
5.8.2	Αποθήκη δεδομένων 58
5.8.3	Τεχνολογία χωρίς διακομιστές..... 58
5.8.4	Ενημέρωση σχετικά με τη διαδικασία 59
5.9	Τεχνολογίες που υποστηρίζουν την ανάλυση σε πραγματικό χρόνο..... 59
5.9.1	Συσκευή αποθήκης δεδομένων 59
5.9.2	Κινητή τεχνολογία..... 60
5.10	Περιοχές εφαρμογής Επιχειρηματικής ευφυΐας 60
6	Κεφάλαιο: «Εργαλεία οπτικοποίησης δεδομένων» 63
6.1	Γενικά στοιχεία 63
6.2	Tableau 63
6.3	Qlikview 64

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

6.4	FusionCharts.....	65
6.5	Highcharts	66
6.6	Datawrapper	66
6.7	Plotly	67
6.8	Sisense.....	68
6.9	Tableau Public.....	68
6.10	Microsoft Power BI.....	69
6.11	Σύγκριση Εργαλείων οπτικοποίησης δεδομένων.....	70
7	Κεφάλαιο: «Συμπεράσματα»	71
	Βιβλιογραφία.....	73
	Πνευματικά δικαιώματα.....	76

Κατάλογος Σχημάτων - Πινάκων

Σχήμα 1.1: Σχηματική απεικόνιση της διαφοράς μεταξύ διερευνητικής και επιβεβαιωτικής ανάλυσης δεδομένων.....	4
Σχήμα 1.2: Διαδικασία ανάλυσης δεδομένων.....	5
Σχήμα 1.3: Οκτώ τύποι ποσοτικών μηνυμάτων, σύμφωνα με τον Stephen Few.....	8
Σχήμα 1.4: Ανάλυση παλινδρόμησης.....	11
Σχήμα 1.5: Ανάλυση και οπτικοποίηση δεδομένων στο ελεύθερο λογισμικό για αριθμητικούς υπολογισμούς DataMelt.....	15
Σχήμα 1.6: Ανάλυση και οπτικοποίηση δεδομένων στο ELKI.....	16
Σχήμα 1.7: Ανάλυση και οπτικοποίηση δεδομένων στην Orange.....	17
Σχήμα 1.8: Ανάλυση και οπτικοποίηση δεδομένων στην Past.....	17
Σχήμα 1.9: Συνδυασμένα δεδομένα LHCb και CMS στο αντικειμενοστρεφές πρόγραμμα ROOT.....	18
Σχήμα 1.10: Απεικόνιση δεδομένων PSD του ECG χρησιμοποιώντας scipy.....	19
Σχήμα 1.11: Περιβάλλον ανάλυσης βιβλιοθήκη λογισμικού pandas.....	19
Σχήμα 2.1: Παράδειγμα διαγράμματος ράβδων.....	28
Σχήμα 2.2: Παράδειγμα «Πολυτροπικού» ιστογράμματος.....	29
Σχήμα 2.3: Παράδειγμα Διάγραμμα διασποράς.....	31
Σχήμα 2.4: Παράδειγμα Προγράμματος έργου.....	32
Σχήμα 2.5: Παράδειγμα streamgraph.....	33
Σχήμα 2.6: Παράδειγμα απεικόνισης πληροφοριών treemapping.....	34
Σχήμα 2.7: Παράδειγμα χάρτη θερμικής κατανομής (heatmap).....	35
Σχήμα 2.8: Αυτό το διάγραμμα Spider αντιπροσωπεύει το διαθέσιμο προϋπολογισμό σε σχέση με τις πραγματικές δαπάνες για έναν συγκεκριμένο οικονομικό οργανισμό.....	36
Σχήμα 3.1: Κριτήρια εκτίμησης Μέτρησης απόδοσης στην διαχείριση έργων.....	48
Σχήμα 5.1: Αρχική σελίδα από το λογισμικό Tableau.....	64
Σχήμα 5.2: Αρχική σελίδα από το λογισμικό Qlikview.....	65
Σχήμα 5.3: Αρχική σελίδα από το λογισμικό fusioncharts.....	65
Σχήμα 5.4: Παράδειγμα δημιουργίας διαδραστικών διαγραμμάτων διαφόρων έργων στο διαδίκτυο με το λογισμικό Highcharts.....	66
Σχήμα 5.5: Παράδειγμα δημιουργίας διαδραστικών χαρτών διαφόρων δεδομένων στο διαδίκτυο με το λογισμικό Datawrapper.....	67
Πίνακας 6.1: Σύγκριση Εργαλείων οπτικοποίησης δεδομένων.....	70

Συντομογραφίες - Απόδοση όρων

ΑΕΠ: Ακαθάριστο Εθνικό Προϊόν

ΑΣΧ: Αναπτυξιακών Στόχων της Χιλιετίας

ΟΗΕ: Οργανισμός Ηνωμένων εθνών

AIDC: Automatic identification and data capture (Αυτόματη αναγνώριση και καταγραφή δεδομένων)

AWS: Amazon Web Services

BAM: Business activity monitoring (Παρακολούθηση επιχειρηματικής δραστηριότητας)

Barcode: γραμμωτοί κώδικες

BI: Business Intelligence (Επιχειρηματική ευφυΐα)

Big data: μεγάλα δεδομένα

BLOB: Binary Large Object (δυναμικό μεγάλο αντικείμενο)

BPM: Business process management (διαχείριση επιχειρησιακών διαδικασιών)

CDA: confirmatory data analysis (ανάλυση επιβεβαιωτικών δεδομένων)

CEP: Complex event processing (σύνθετη επεξεργασία συμβάντων)

CERN: Conseil Européen pour la Recherche Nucléaire (Ευρωπαϊκό Συμβούλιο για την Πυρηνική Έρευνα)

Competitive intelligence: ανταγωνιστική νοημοσύνη

CRM: Customer Relationship Management (Διαχείριση Πελατειακών Σχέσεων)

CSV: Comma Separated Values (Τιμές διαχωρισμένες με κόμματα)

Data warehouse: Βάση δεδομένων

DBA: Database administrator (διαχειριστής βάσεων δεδομένων)

DBMS: database management systems (συστήματα διαχείρισης βάσεων δεδομένων)

EAI: Enterprise application integration (λειτουργίες ολοκλήρωσης εφαρμογών για επιχειρήσεις)

EDA: exploratory data analysis (διερευνητική ανάλυση δεδομένων)

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

ESP: Event stream processing (Επεξεργασία ροής συμβάντων)

ETL: Extraction, Transformation, Loading (Εξόρυξη, μετασχηματισμός, φόρτωση)

Heatmap: χάρτες θερμικής κατανομής

IoT: The Internet of Things (Διαδίκτυο των πραγμάτων)

KDD: knowledge discovery in databases (αποκάλυψη γνώσεων σε βάσεις δεδομένων / εξόρυξη δεδομένων)

KNIME: Konstanz Information Miner (βάση πληροφόρησης Konstanz)

KPIs: key performance indicators (βασικοί δείκτες απόδοσης)

MBI: Mobile Business Intelligence (Επιχειρηματική ευφυΐα κινητού)

MDA: multidimensional analysis (πολυδιάστατης ανάλυσης)

MSSO: Multiple Source Simple Output (απλή έξοδος πολλαπλών πηγών)

NCA: necessary analysis of conditions (αναγκαία ανάλυση συνθηκών)

OLAP: Online analytical processing (Ηλεκτρονική αναλυτική επεξεργασία)

RCA: Root cause analysis (Ανάλυση των βαθύτερων αιτίων)

RFID: Radio-frequency identification (αναγνώριση ραδιοσυχνοτήτων)

RTBI: Real-time business intelligence (Επιχειρηματική νοημοσύνη σε πραγματικό χρόνο)

SAS: Statistical Analysis System (Σύστημα στατιστικής ανάλυσης)

SOFA: Statistics Open For All (Στατιστικά Ανοιχτά για Όλους)

TWAP: time-weighted average price (μέση σταθμισμένη τιμή χρόνου)

VWAP: Volume Weighted Average Price (Μέση σταθμισμένη τιμή όγκου)

UNICEF: United Nations International Children's Emergency Fund (Διεθνές Ταμείο Επείγουσας Βοήθειας των Ηνωμένων Εθνών για τα Παιδιά)

WSN: Wireless sensor network (ασύρματο δίκτυο αισθητήρων)

1 Κεφάλαιο: «Εισαγωγή»

Το θέμα της παρούσης πτυχιακής εργασίας είναι: «Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας» και αναλύεται ως εξής:

Στο 2^ο κεφάλαιο: «Ανάλυση δεδομένων» προτάσσονται γενικά στοιχεία του θέματος, η διαδικασία ανάλυσης δεδομένων (απαιτήσεις – συλλογή - επεξεργασία - καθαρισμός δεδομένων, διερευνητική ανάλυση δεδομένων, μοντελοποίηση και αλγόριθμοι), τα ποσοτικά μηνύματα, οι τεχνικές ανάλυσης ποσοτικών δεδομένων, τα προβλήματα αποτελεσματικής ανάλυσης (γνωσιακές προκαταλήψεις, αναξιοπιστία) και τέλος τα ελεύθερα λογισμικά για την ανάλυση δεδομένων.

Στο 3^ο κεφάλαιο: «Οπτικοποίηση δεδομένων» παρατίθενται γενικά στοιχεία, διάφοροι ορισμοί / έννοιες, τα χαρακτηριστικά των αποτελεσματικών γραφικών απεικονίσεων, τα ποσοτικά μηνύματα, η οπτική αντίληψη και απεικόνιση δεδομένων, η ιστορική αναδρομή της απεικόνισης δεδομένων και διάφορα παραδείγματα διαγραμμάτων χρήσης οπτικοποίησης δεδομένων (ραβδόγραμμα, ιστόγραμμα, διάγραμμα διασποράς, διάγραμμα έργου, Streamgraph, Treemapping, Heatmap, διαγράμματα Spider), τέλος η αρχιτεκτονική παρουσίασης δεδομένων.

Στο 4^ο κεφάλαιο: «Επιχειρηματική ευφυΐα» αναλύονται οι λειτουργίες και τεχνολογίες της επιχειρηματικής ευφυΐας, τα χαρακτηριστικά της επιχειρηματικής ευφυΐας (σύγκριση με την ανταγωνιστική νοημοσύνη και τις επιχειρηματικές αναλύσεις), τα στοιχεία της επιχειρηματικής ευφυΐας, τα δεδομένα επιχειρηματικών λειτουργιών (μη δομημένα δεδομένα έναντι ημιδομημένων δεδομένων, περιορισμοί ημιδομημένων και αδόμητων δεδομένων), τα μεταδεδομένα και οι εφαρμογές επιχειρηματικής ευφυΐας (μετρήσεις απόδοσης, συγκριτική αξιολόγηση, διαχείριση γνώσης).

Στο 5^ο κεφάλαιο: «Επιχειρηματική ευφυΐα σε πραγματικό χρόνο» στοιχειοθετείται ο σκοπός της επιχειρηματικής ευφυΐας, τα διάφορα στοιχεία επιχειρηματικής ευφυΐας, τα τεχνολογικά στοιχεία επιχειρηματικής ευφυΐας, η επιχειρηματική νοημοσύνη σε πραγματικό χρόνο, η λήψη αποφάσεων επιχειρηματικής ευφυΐας, η καθυστέρηση της επιχειρηματικής ευφυΐας, οι αρχιτεκτονικές επιχειρηματικής ευφυΐας (βάση συμβάντων αποθήκη δεδομένων, τεχνολογία χωρίς διακομιστές, ενημέρωση σχετικά με τη διαδικασία) και οι τεχνολογίες που υποστηρίζουν την ανάλυση σε πραγματικό χρόνο (συσκευή αποθήκης δεδομένων, κινητή τεχνολογία), περιοχές εφαρμογής επιχειρηματικής ευφυΐας.

Στο 6^ο κεφάλαιο: «Εργαλεία οπτικοποίησης δεδομένων» καταγράφονται γενικά στοιχεία για τα εργαλεία οπτικοποίησης δεδομένων, δηλαδή για τα: Tableau, Qlikview, FusionCharts, Highcharts, Datawrapper, Plotly, Sisense, Tableau Public, Microsoft Power BI.

Στο 7^ο και τελευταίο κεφάλαιο τα: «Συμπεράσματα» παρουσιάζονται βασικά στοιχεία για τα εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας, τα οποία προέκυψαν από την βιβλιογραφική έρευνα που προηγήθηκε.

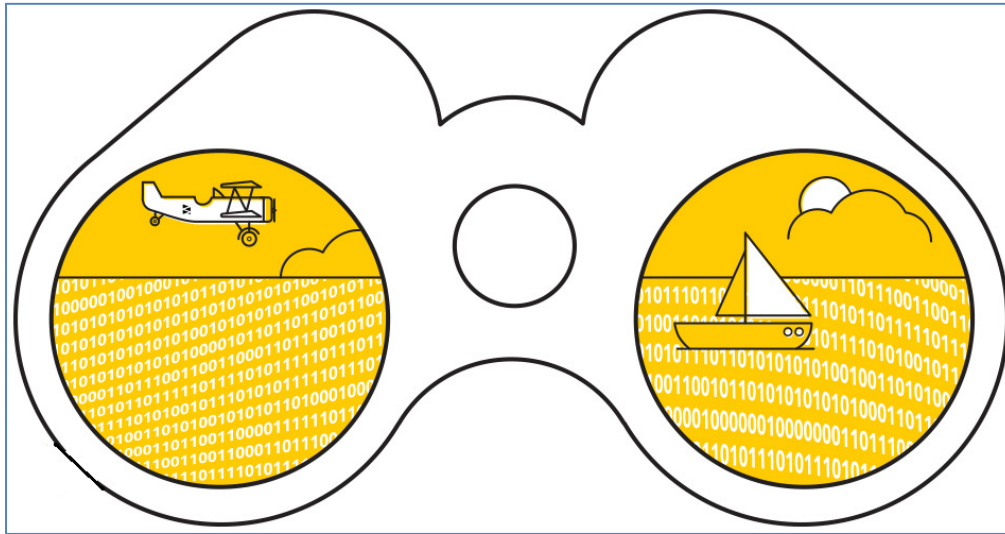
2 Κεφάλαιο: «Ανάλυση δεδομένων»

2.1 Γενικά στοιχεία

Η ανάλυση δεδομένων είναι μια διαδικασία επιθεώρησης, καθαρισμού (πρόκειται για την διεργασία ανίχνευσης και διόρθωσης κατεστραμμένων ή ανακριβών αρχείων από μία βάση δεδομένων), μετασχηματισμού (πρόκειται για την διεργασία μετατροπής στοιχείων / δεδομένων από μία μορφή ή δομή σε άλλη) και μοντελοποίησης (πρόκειται για την διεργασία σχεδιασμού ενός μοντέλου δεδομένων για ένα σύστημα πληροφοριών με την χρήση συγκεκριμένων έγκριτων τεχνικών) δεδομένων με στόχο την ανακάλυψη χρήσιμων πληροφοριών, την ενημέρωση συμπερασμάτων και τη λήψη αποφάσεων. Η ανάλυση δεδομένων έχει πολλαπλές πτυχές και προσεγγίσεις, που περιλαμβάνουν διαφορετικές τεχνικές με ποικίλα ονόματα, ενώ χρησιμοποιούνται σε διαφορετικούς τομείς των επιχειρήσεων, της επιστήμης και της κοινωνικής επιστήμης. Στη σημερινή επιχειρηματική δραστηριότητα, η ανάλυση δεδομένων παίζει ρόλο στη λήψη αποφάσεων πιο επιστημονικής φύσης και στη διευκόλυνση της αποτελεσματικής λειτουργίας της επιχείρησης (Παπαδημητρίου, Η Ανάλυση Δεδομένων-Παραγοντική Ανάλυση των Αντιστοιχιών - Ιεραρχική Ταξινόμηση και άλλες Μέθοδοι, 2007).

Η εξόρυξη δεδομένων είναι μια συγκεκριμένη τεχνική ανάλυσης δεδομένων που επικεντρώνεται στη μοντελοποίηση και την ανακάλυψη γνώσεων για λόγους πρόβλεψης και όχι για καθαρά περιγραφικούς σκοπούς (αυτή η διεργασία περιλαμβάνει μεθόδους της μηχανικής μάθησης, των στατιστικών αλλά και των υπολογιστικών συστημάτων βάσεων δεδομένων). Αντίθετα η επιχειρησιακή ευφυΐα καλύπτει ανάλυση δεδομένων που εξαρτάται σε μεγάλο βαθμό από τη συγκέντρωση, εστιάζοντας κυρίως στις επιχειρηματικές πληροφορίες (ακόμη περιέχει τις στρατηγικές και τις τεχνολογίες που χρησιμοποιούνται από τις εκάστοτε επιχειρήσεις για την ανάλυση δεδομένων επιχειρηματικών πληροφοριών). Σε στατιστικές εφαρμογές, η ανάλυση δεδομένων μπορεί να χωριστεί σε περιγραφικά στατιστικά στοιχεία, διερευνητική ανάλυση δεδομένων (EDA) (είναι μια προσέγγιση που αναλύει σύνολα δεδομένων για να συνοψίσει τα κύρια χαρακτηριστικά τους, συχνά με οπτικές μεθόδους) και ανάλυση επιβεβαιωτικών δεδομένων (confirmatory data analysis - CDA) (πρόκειται για υπόθεση που μπορεί να ελεγχθεί με βάση την παρατήρηση μιας διαδικασίας που διαμορφώνεται μέσω ενός συνόλου τυχαίων μεταβλητών). Η EDA επικεντρώνεται στην

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας ανακάλυψη νέων χαρακτηριστικών στα δεδομένα ενώ η CDA επικεντρώνεται στην επιβεβαίωση ή παραποίηση των υφιστάμενων υποθέσεων(Blitz, 2017).



Σχήμα 2.1: Σχηματική απεικόνιση της διαφοράς μεταξύ διερευνητικής και επιβεβαιωτικής ανάλυσης δεδομένων.

Πηγή: (Blitz, 2017).

Η Προγνωστική ανάλυση επικεντρώνεται στην εφαρμογή στατιστικών μοντέλων για πρόβλεψη ή ταξινόμηση, ενώ οι αναλύσεις κειμένου εφαρμόζουν στατιστικές, γλωσσικές και δομικές τεχνικές για την εξαγωγή και ταξινόμηση πληροφοριών από κειμενικές πηγές, ένα είδος αδόμητων δεδομένων. Όλα τα παραπάνω είναι ποικιλίες ανάλυσης δεδομένων(Blitz, 2017).

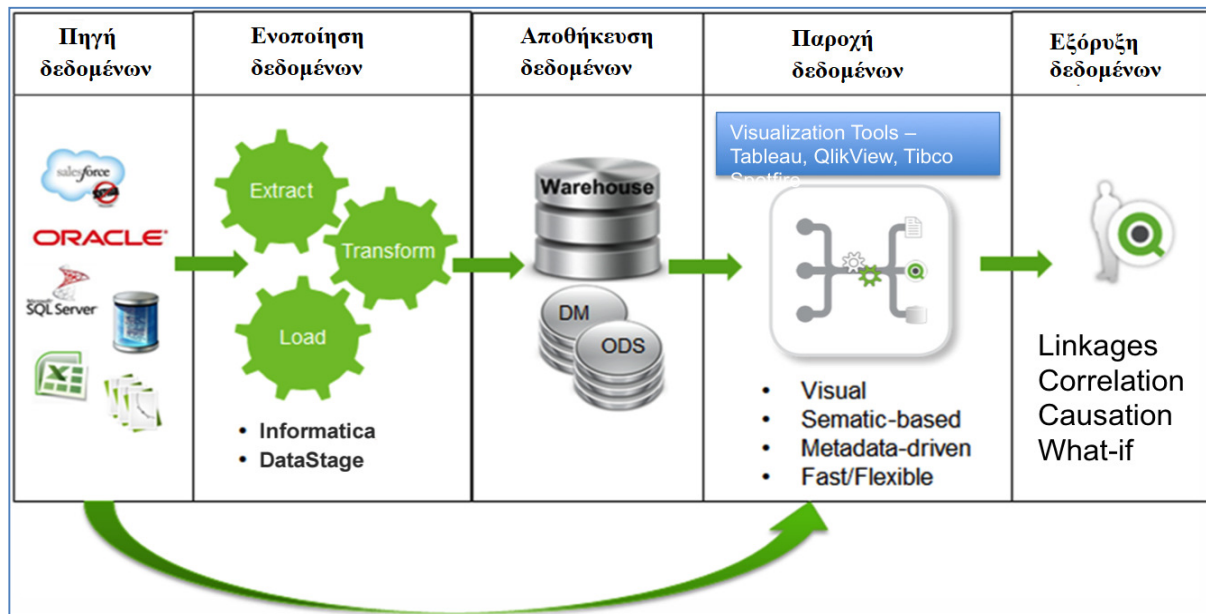
Η ενσωμάτωση δεδομένων αποτελεί πρόδρομο της ανάλυσης δεδομένων και η ανάλυση των δεδομένων είναι στενά συνδεδεμένη στην οπτικοποίηση δεδομένων και τη διάδοση των δεδομένων, επιπλέον συνδυάζει στοιχεία / δεδομένα που διέρχονται από διαφορετικές πηγές και παρέχονται στους χρήστες ως μια ενιαία οντότητα. Ο όρος ανάλυση δεδομένων χρησιμοποιείται μερικές φορές ως συνώνυμο για τη μοντελοποίηση δεδομένων.

2.2 Διαδικασία ανάλυσης δεδομένων

Η ανάλυση αναφέρεται στο διαχωρισμό ενός συνόλου στα ξεχωριστά συστατικά του για ατομική εξέταση. Η ανάλυση δεδομένων είναι μια διαδικασία για την απόκτηση ανεπεξέργαστων δεδομένων και τη μετατροπή τους σε πληροφορίες χρήσιμες για τη λήψη αποφάσεων από τους χρήστες. Τα δεδομένα συλλέγονται και αναλύονται για να απαντούν σε ερωτήσεις, να δοκιμάζουν υποθέσεις ή να διαψεύδουν θεωρίες(Παπαδημητρίου, Η Ανάλυση

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

Δεδομένων-Παραγοντική Ανάλυση των Αντιστοιχιών - Ιεραρχική Ταξινόμηση και άλλες Μέθοδοι, 2007).



Σχήμα 2.2: Διαδικασία ανάλυσης δεδομένων.

Πηγή δεδομένων (Data Source), Ενοποίηση δεδομένων (Data Consolidation), Αποθήκευση δεδομένων (Data Storage), Παροχή δεδομένων (Data Provisioning) και Εξόρυξη δεδομένων (Data Discovery).

Πηγή: (Business Analytics, 2018).

Ο μαθηματικός στατιστικολόγος John Tukey (1961) όρισε την ανάλυση δεδομένων ως εξής: «Διαδικασίες για την ανάλυση δεδομένων, τεχνικές για την ερμηνεία των αποτελεσμάτων τέτοιων διαδικασιών, τρόπους σχεδιασμού της συλλογής δεδομένων για να γίνει η ανάλυση ευκολότερη, ακριβέστερη»(Business Analytics, 2018).

Υπάρχουν διάφορες φάσεις που μπορούν να διακριθούν, που περιγράφονται παρακάτω. Οι φάσεις είναι επαναληπτικές, καθώς η ανατροφοδότηση από τις μεταγενέστερες φάσεις μπορεί να οδηγήσει σε πρόσθετη εργασία σε προηγούμενες φάσεις.

2.2.1 Απαιτήσεις δεδομένων

Τα δεδομένα είναι απαραίτητα ως εισροές στην ανάλυση, η οποία καθορίζεται με βάση τις απαιτήσεις εκείνων που κατευθύνουν την ανάλυση ή τους πελάτες (που θα χρησιμοποιήσουν το τελικό προϊόν της ανάλυσης). Ο γενικός τύπος της οντότητας επί της οποίας θα συλλέγονται τα δεδομένα αναφέρεται ως πειραματική μονάδα (για παράδειγμα ένα άτομο ή πληθυσμός ανθρώπων). Μπορούν να προσδιοριστούν και να ληφθούν συγκεκριμένες μεταβλητές που αφορούν έναν πληθυσμό (για παράδειγμα ηλικία και εισόδημα). Τα δεδομένα

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας μπορεί να είναι αριθμητικά ή κατηγορικά (δηλαδή, μια ετικέτα κειμένου για αριθμούς)(Business Analytics, 2018).

2.2.2 Συλλογή δεδομένων

Τα δεδομένα συλλέγονται από διάφορες πηγές. Οι απαιτήσεις μπορούν να κοινοποιούνται από τους αναλυτές στους θεματοφύλακες των δεδομένων, όπως το προσωπικό της τεχνολογίας, της πληροφορίας εντός ενός οργανισμού. Τα δεδομένα μπορούν επίσης να συλλέγονται, όπως κάμερες κυκλοφορίας, δορυφόροι, συσκευές εγγραφής κλπ. Μπορεί επίσης να ληφθεί μέσω συνεντεύξεων, λήψεων από πηγές στο διαδίκτυο ή ανάγνωσης τεκμηρίωσης(Business Analytics, 2018).

2.2.3 Επεξεργασία δεδομένων

Τα δεδομένα που ελήφθησαν αρχικά πρέπει να υποβληθούν σε επεξεργασία ή να οργανωθούν για ανάλυση. Για παράδειγμα, αυτά μπορεί να περιλαμβάνουν την τοποθέτηση δεδομένων σε σειρές και στήλες σε μορφή πίνακα (δηλαδή, Δομημένα δεδομένα) για περαιτέρω ανάλυση, όπως σε ένα υπολογιστικό φύλλο ή σε ένα στατιστικό λογισμικό(Business Analytics, 2018).

2.2.4 Καθαρισμός δεδομένων

Αφού γίνει η επεξεργασία και η οργάνωση, τα δεδομένα ενδέχεται να είναι ελλιπή, να περιέχουν αντίγραφα ή να περιέχουν σφάλματα. Η ανάγκη για τον καθαρισμό των δεδομένων θα προκύψει από προβλήματα σχετικά με τον τρόπο εισαγωγής και αποθήκευσης δεδομένων. Ο καθαρισμός δεδομένων είναι η διαδικασία πρόληψης και διόρθωσης αυτών των σφαλμάτων. Οι συνηθισμένες εργασίες περιλαμβάνουν την αντιστοίχιση αρχείων, τον εντοπισμό ανακρίβειας των δεδομένων, τη συνολική ποιότητα των υπαρχόντων δεδομένων και την κατάτμηση της στήλης(Business Analytics, 2018).

Τέτοια προβλήματα δεδομένων μπορούν επίσης να εντοπιστούν μέσω μιας ποικιλίας αναλυτικών τεχνικών. Για παράδειγμα, με τα οικονομικά στοιχεία, τα σύνολα για συγκεκριμένες μεταβλητές μπορούν να συγκριθούν με ξεχωριστά δημοσιευμένους αριθμούς που θεωρούνται αξιόπιστοι. Μπορούν επίσης να αναθεωρηθούν ασυνήθιστα ποσά πάνω ή κάτω από τα προκαθορισμένα όρια. Υπάρχουν διάφοροι τύποι καθαρισμού δεδομένων που εξαρτώνται από τον τύπο δεδομένων όπως αριθμούς τηλεφώνου, διευθύνσεις ηλεκτρονικού ταχυδρομείου, εργοδότες κ.λπ. Οι μέθοδοι ποσοτικών δεδομένων για την ανίχνευση των εξωστρεφών μπορούν να χρησιμοποιηθούν για την απαλλαγή από πιθανά λανθασμένα

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας καταχωρημένα δεδομένα. Οι ορθογραφικοί έλεγχοι των κειμενικών στοιχείων μπορούν να χρησιμοποιηθούν για να μειώσουν την ποσότητα των λέξεων με σφάλματα (Κύρκος, 2015).

2.2.5 Διερευνητική ανάλυση δεδομένων

Μόλις τα δεδομένα καθαριστούν, μπορούν να αναλυθούν. Οι αναλυτές μπορούν να εφαρμόσουν μια ποικιλία τεχνικών που αναφέρονται ως διερευνητική ανάλυση δεδομένων για να ξεκινήσουν την κατανόηση των μηνυμάτων που περιέχονται στα δεδομένα. Η διαδικασία εξερεύνησης μπορεί να οδηγήσει σε πρόσθετο καθαρισμό δεδομένων ή πρόσθετα αιτήματα για δεδομένα, επομένως οι δραστηριότητες αυτές ενδέχεται να έχουν επαναληπτικό χαρακτήρα. Μπορούν να δημιουργηθούν περιγραφικά στατιστικά στοιχεία, όπως ο μέσος όρος ή ο διάμεσος, για την κατανόηση των δεδομένων. Η οπτικοποίηση δεδομένων μπορεί επίσης να χρησιμοποιηθεί για την εξέταση των δεδομένων σε γραφική μορφή, για να αποκτηθεί πρόσθετη γνώση σχετικά με τα μηνύματα μέσα στα δεδομένα (Κύρκος, 2015).

2.2.6 Μοντελοποίηση και αλγόριθμοι

Μαθηματικοί τύποι ή μοντέλα που ονομάζονται αλγόριθμοι μπορούν να εφαρμοστούν στα δεδομένα για τον προσδιορισμό των σχέσεων μεταξύ των μεταβλητών, όπως συσχετισμός ή αιτιώδης συνάφεια. Σε γενικές γραμμές, μπορούν να αναπτυχθούν μοντέλα για την αξιολόγηση μιας συγκεκριμένης μεταβλητής στα δεδομένα που βασίζονται σε άλλες μεταβλητές στα δεδομένα, με κάποια υπολειμματικά σφάλματα ανάλογα με την ακρίβεια του μοντέλου, για παράδειγμα (Business Analytics, 2018):

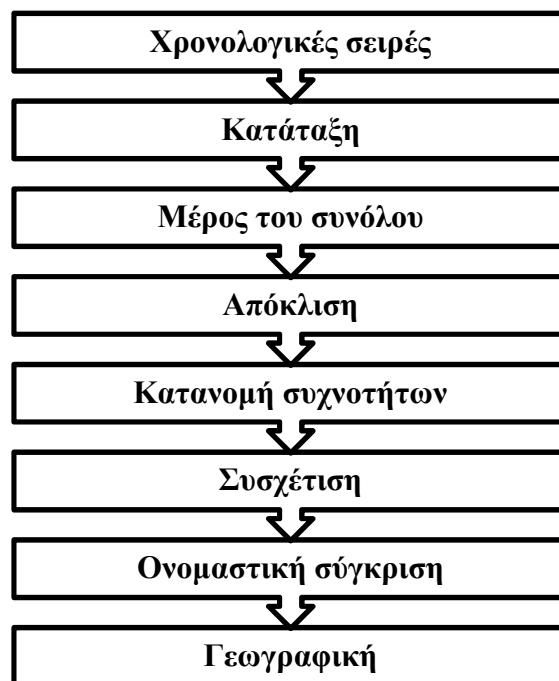
$$\text{Δεδομένα} = \text{Μοντέλο} + \text{Σφάλμα}$$

Οι στατιστικές παρεμβάσεων περιλαμβάνουν τεχνικές μέτρησης σχέσεων μεταξύ συγκεκριμένων μεταβλητών. Για παράδειγμα, η ανάλυση παλινδρόμησης μπορεί να χρησιμοποιηθεί για να μοντελοποιήσει εάν μια αλλαγή στη διαφήμιση (ανεξάρτητη μεταβλητή X) εξηγεί τη μεταβολή των πωλήσεων (εξαρτημένη μεταβλητή Y). Από μαθηματικούς όρους, το Y (πωλήσεις) είναι συνάρτηση του X (διαφήμιση). Μπορεί να περιγραφεί ως σφάλμα: $Y = aX + b$,

όπου το μοντέλο έχει σχεδιαστεί έτσι ώστε τα a και b να ελαχιστοποιούν το σφάλμα όταν το μοντέλο προβλέπει Y για ένα δεδομένο εύρος τιμών του X . Οι αναλυτές μπορούν να επιχειρήσουν να δημιουργήσουν μοντέλα που είναι περιγραφικά τα δεδομένα για την απλούστευση της ανάλυσης και την επικοινωνία των αποτελεσμάτων.

2.3 Ποσοτικά μηνύματα

Ο Stephen Few (αναλυτής στον τομέα της απεικόνισης δεδομένων) περιγράφει οκτώ τύπους ποσοτικών μηνυμάτων που οι χρήστες ενδέχεται να επιχειρήσουν να κατανοήσουν ή να επικοινωνήσουν από ένα σύνολο δεδομένων και τα σχετικά γραφήματα που χρησιμοποιούνται για να βοηθήσουν στην επικοινωνία του μηνύματος (Σχήμα 1.1). Οι πελάτες που καθορίζουν τις απαιτήσεις και οι αναλυτές που πραγματοποιούν την ανάλυση δεδομένων μπορούν να εξετάσουν αυτά τα μηνύματα κατά τη διάρκεια της διαδικασίας (Κύρκος, 2015).



Σχήμα 2.3: Οκτώ τύποι ποσοτικών μηνυμάτων, σύμφωνα με τον Stephen Few.

Πηγή: (Κύρκος, 2015).

Αναλυτικά οι οκτώ τύποι ποσοτικών μηνυμάτων, σύμφωνα με τον Stephen Few (Κύρκος, 2015):

Χρονολογικές σειρές: Μια ενιαία μεταβλητή καταγράφεται σε μια χρονική περίοδο, όπως το ποσοστό ανεργίας για μια δεκαετή περίοδο. Μπορεί να χρησιμοποιηθεί ένα γράφημα γραμμής για να αποδειχθεί η τάση.

Κατάταξη: Οι κατηγοριοποιημένες υποδιαίρεσεις κατατάσσονται σε αύξουσα ή φθίνουσα σειρά, όπως η κατάταξη των επιδόσεων των πωλήσεων (το μέτρο) από τα άτομα πωλήσεων (κατηγορία, με κάθε πωλητή κατηγοριοποιημένη υποδιαίρεση) κατά τη διάρκεια

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας μιας μόνο περιόδου. Ένα διάγραμμα ράβδων μπορεί να χρησιμοποιηθεί για να δείξει τη σύγκριση μεταξύ των πωλητών.

Μέρος του συνόλου: Οι κατηγοριοποιημένες υποδιαιρέσεις μετρούνται σε σχέση με το σύνολο (δηλαδή ένα ποσοστό από το 100%). Ένα διάγραμμα πίτας ή ένα διάγραμμα ράβδων μπορεί να δείξει τη σύγκριση των λόγων, όπως το μερίδιο αγοράς που αντιπροσωπεύουν οι ανταγωνιστές σε μια αγορά.

Απόκλιση: Οι κατηγοριοποιημένες υποδιαιρέσεις συγκρίνονται με μια αναφορά, όπως η σύγκριση των πραγματικών έναντι των δαπανών του προϋπολογισμού για διάφορα τμήματα μιας επιχείρησης για μια δεδομένη χρονική περίοδο. Ένα διάγραμμα ράβδων μπορεί να δείξει σύγκριση του πραγματικού με το ποσό αναφοράς.

Κατανομή συχνοτήτων: Δείχνει τον αριθμό των παρατηρήσεων μιας συγκεκριμένης μεταβλητής για δεδομένο διάστημα, όπως ο αριθμός των ετών κατά τα οποία η χρηματιστηριακή απόδοση είναι μεταξύ διαστημάτων όπως 0-10%, 11-20% κλπ. Ένα ιστόγραμμα, ένας τύπος του ραβδογράμματος, μπορεί να χρησιμοποιηθεί για την ανάλυση αυτή.

Συσχέτιση: Σύγκριση μεταξύ παρατηρήσεων που αντιπροσωπεύονται από δύο μεταβλητές (X, Y) για να προσδιοριστεί εάν τείνουν να κινούνται στις ίδιες ή αντίθετες κατευθύνσεις. Για παράδειγμα, σχεδιάζεται η ανεργία (X) και ο πληθωρισμός (Y) για ένα δείγμα μηνών. Για αυτό το μήνυμα χρησιμοποιείται συνήθως μια γραφική παράσταση σκέδασης.

Ονομαστική σύγκριση: Συγκρίνοντας κατηγορηματικές υποδιαιρέσεις σε καμία συγκεκριμένη σειρά, όπως τον όγκο πωλήσεων ανά κωδικό προϊόντος. Για τη σύγκριση αυτή μπορεί να χρησιμοποιηθεί διάγραμμα ράβδων.

Γεωγραφική ή γεωχωρική: Σύγκριση μεταβλητής σε χάρτη ή διάταξη, όπως το ποσοστό ανεργίας ανά κράτος ή ο αριθμός των ατόμων στους διάφορους ορόφους ενός κτιρίου. Ένα χαρτογράφημα είναι ένα τυπικό γραφικό που χρησιμοποιείται.

2.4 Τεχνικές ανάλυσης ποσοτικών δεδομένων

Ο Jonathan Koomey έχει συστήσει μια σειρά βέλτιστων πρακτικών για την κατανόηση ποσοτικών δεδομένων. Αυτά περιλαμβάνουν (Business Analytics, 2018):

- Έλεγχο των ακατέργαστων δεδομένων για ανωμαλίες πριν εκτελεστεί η ανάλυση.

- Εκ νέου εκτέλεση σημαντικών υπολογισμών, όπως η επαλήθευση των στηλών των δεδομένων που οδηγούνται από τον τύπο.
- Επιβεβαίωση ότι τα βασικά σύνολα είναι το άθροισμα των υποσυνόλων.
- Έλεγχο των σχέσεων μεταξύ των αριθμών που θα πρέπει να σχετίζονται με έναν προβλέψιμο τρόπο, όπως αναλογίες με την πάροδο του χρόνου.
- Κανονικοποίηση των αριθμών για τη διευκόλυνση των συγκρίσεων, όπως ανάλυση των ποσών ανά άτομο ή σε σχέση με το ΑΕΠ ή ως τιμή δείκτη σε σχέση με ένα έτος βάσης.
- Αντιμετώπιση των προβλημάτων σε συστατικά τμήματα αναλύοντας τους παράγοντες που οδήγησαν στα αποτελέσματα, όπως η ανάλυση της απόδοσης της «καθαρής θέσης» (δηλαδή της διαφοράς του συνολικού ενεργητικού).

Για τις υπό εξέταση μεταβλητές, οι αναλυτές λαμβάνουν χαρακτηριστικά περιγραφικά στατιστικά στοιχεία για αυτά, όπως η μέση, η διάμεση και η τυπική απόκλιση. Μπορούν επίσης να αναλύσουν τη διανομή των βασικών μεταβλητών για να δουν τον τρόπο με τον οποίο η ατομική ομάδα συγκεντρώνεται γύρω από το μέσο όρο.

Οι αναλυτές μπορούν να χρησιμοποιούν αξιόπιστες στατιστικές μετρήσεις για την επίλυση ορισμένων αναλυτικών προβλημάτων. Ο έλεγχος της υπόθεσης χρησιμοποιείται όταν μια συγκεκριμένη υπόθεση σχετικά με την πραγματική κατάσταση των πραγμάτων γίνεται από τον αναλυτή και συλλέγονται δεδομένα για να καθοριστεί εάν η κατάσταση αυτή είναι αληθινή ή ψευδής (βέβαια μπορεί να ελεγχθεί με γνώμονα την εξέταση μιας διεργασίας που ανασχηματίζεται μέσω ενός συνόλου τυχαίων μεταβλητών). Για παράδειγμα, η υπόθεση μπορεί να είναι ότι η «Ανεργία δεν έχει επίδραση στον πληθωρισμό», η οποία σχετίζεται με μια οικονομική έννοια που ονομάζεται καμπύλη Phillips¹. Ο έλεγχος της υπόθεσης περιλαμβάνει την εξέταση της πιθανότητας σφαλμάτων τύπου I και τύπου II², τα οποία

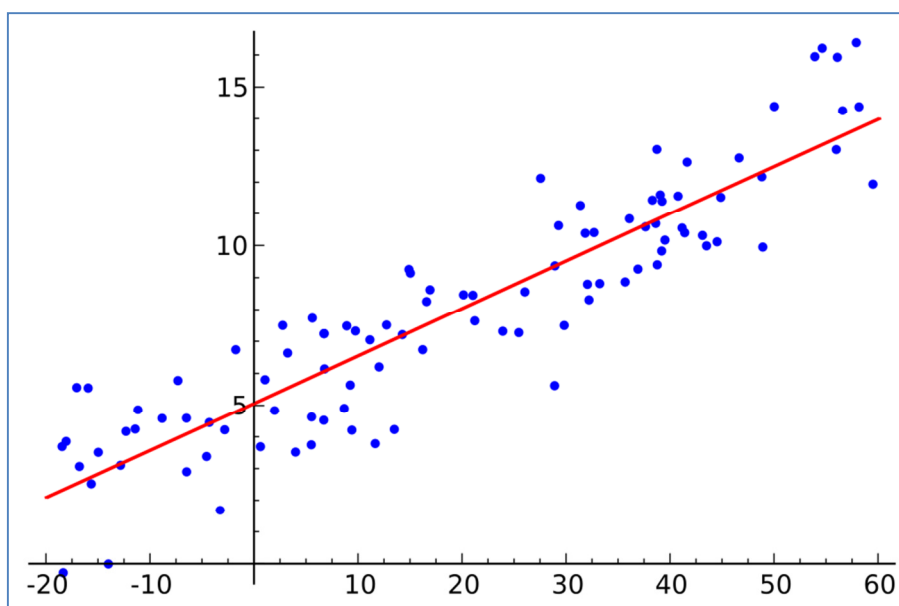
¹ Η καμπύλη Phillips είναι ένα οικονομετρικό μοντέλο μιας εξίσωσης, το οποίο κατονομάζεται από τον William Phillips, περιγράφοντας μια ιστορική αντίστροφη σχέση μεταξύ των ποσοστών ανεργίας και των αντίστοιχων ποσοστών αύξησης των μισθών που προκύπτουν από μια οικονομία. Πιο απλά, η μείωση της ανεργίας σε μια οικονομία θα συσχετιστεί με υψηλότερα ποσοστά αύξησης των μισθών. Ο Phillips δεν δήλωσε ο ίδιος ότι υπάρχει σχέση μεταξύ απασχόλησης και πληθωρισμού, αυτή η έννοια ήταν μια ασήμαντη παρακμή από τα στατιστικά του ευρήματα. Οι Samuelson και Solow έκαναν τη σύνδεση ρητή και στη συνέχεια ο Milton Friedman από το 1967 έβαλε τη θεωρητική δομή στη θέση του. Με τον τρόπο αυτό, ο Friedman πρότεινε με επιτυχία την επικείμενη κατάρρευση της α-θεωρητικής συσχέτισης του Phillips.

² Στη δοκιμασία στατιστικών υποθέσεων, ένα σφάλμα τύπου I είναι η απόρριψη μιας πραγματικής μηδενικής υπόθεσης, ενώ ένα σφάλμα τύπου II αποτυγχάνει να απορρίψει μια ψευδή μηδενική υπόθεση. Πιο απλά, ένα

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

σχετίζονται με το κατά πόσον τα δεδομένα υποστηρίζουν την αποδοχή ή την απόρριψη της υπόθεσης.

Η ανάλυση παλινδρόμησης (σύνολο στατιστικών διεργασιών για την εκτίμηση των σχέσεων μεταξύ των μεταβλητών) μπορεί να χρησιμοποιηθεί όταν ο αναλυτής προσπαθεί να προσδιορίσει την έκταση στην οποία η ανεξάρτητη μεταβλητή X επηρεάζει τη εξαρτημένη μεταβλητή Y (για παράδειγμα «Σε ποιο βαθμό επηρεάζουν τις μεταβολές του ποσοστού ανεργίας (X) τον ρυθμό πληθωρισμού (Y)»). Αυτή είναι μια προσπάθεια να μοντελοποιηθεί ή να προσαρμοστεί μια γραμμή εξίσωσης ή καμπύλη στα δεδομένα, έτσι ώστε το Y να είναι συνάρτηση του X (Business Analytics, 2018).



Σχήμα 2.4: Ανάλυση παλινδρόμησης.

Τα τυχαία σημεία δεδομένων και η γραμμική τους παλινδρόμηση.

Η απαραίτητη ανάλυση της κατάστασης (Necessary Condition Analysis - NCA) (τεχνική ανάλυσης δεδομένων για τον προσδιορισμό των απαραίτητων συνθηκών στα σύνολα δεδομένων) μπορεί να χρησιμοποιηθεί όταν ο αναλυτής προσπαθεί να καθορίσει την έκταση στην οποία η ανεξάρτητη μεταβλητή X επιτρέπει την μεταβλητή Y (για παράδειγμα «Σε ποιο βαθμό είναι απαραίτητο ένα ορισμένο ποσοστό ανεργίας (X)»). Ενώ η ανάλυση (πολλαπλής) παλινδρόμησης χρησιμοποιεί λογική προσθετικών όπου κάθε μεταβλητή X μπορεί να παράγει το αποτέλεσμα και τα X μπορούν να αντισταθίσουν το ένα το άλλο (είναι επαρκή αλλά όχι απαραίτητα), η αναγκαία ανάλυση συνθηκών (NCA) χρησιμοποιεί λογική

σφάλμα τύπου I είναι να υποθεθεί ψευδώς η ύπαρξη κάτι που δεν υπάρχει εκεί, ενώ ένα σφάλμα τύπου II είναι να υποθεθεί ψευδώς η απουσία κάτι που υπάρχει.

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας αναγκαιότητας, όπου μία ή περισσότερες X - οι μεταβλητές επιτρέπουν την έκβαση του αποτελέσματος, αλλά δεν μπορούν να το παράγουν (είναι απαραίτητες αλλά όχι επαρκείς). Κάθε απαραίτητη προϋπόθεση πρέπει να υπάρχει και η αντιστάθμιση δεν είναι δυνατή (Business Analytics, 2018).

2.5 Προβλήματα αποτελεσματικής ανάλυσης

Τα εμπόδια στην αποτελεσματική ανάλυση μπορεί να υπάρχουν μεταξύ των αναλυτών που πραγματοποιούν την ανάλυση των δεδομένων ή μεταξύ των ακροατών. Διακρίνοντας το γεγονός από τη γνώμη, οι γνωστικές προκαταλήψεις και η αταξία είναι όλες οι προκλήσεις για την ορθή ανάλυση των δεδομένων.

2.5.1 Γεγονός που προκαλεί σύγχυση

Η αποτελεσματική ανάλυση απαιτεί την απόκτηση σχετικών στοιχείων για την απάντηση σε ερωτήσεις, την υποστήριξη συμπεράσματος ή επίσημης γνώμης ή για υποθέσεις δοκιμών. Τα γεγονότα εξ ορισμού είναι αδιάψευστα, που σημαίνει ότι κάθε άτομο που συμμετέχει στην ανάλυση θα πρέπει να είναι σε θέση να τα συμφωνήσει.

Για παράδειγμα, ο ελεγκτής ενός δημόσιου οργανισμού πρέπει να καταλήξει σε μια επίσημη γνώμη σχετικά με το εάν οι οικονομικές καταστάσεις των εταιρειών που διαπραγματεύονται στο χρηματιστήριο είναι «δίκαια δηλωμένες από κάθε ουσιώδη άποψη». Αυτό απαιτεί εκτενή ανάλυση των πραγματικών στοιχείων και αποδεικτικών στοιχείων που υποστηρίζουν τη γνώμη τους. Όταν γίνεται το άλμα από τα γεγονότα στις απόψεις, υπάρχει πάντοτε η πιθανότητα η γνώμη να είναι λανθασμένη (Παπαδημητρίου, Η ανάλυση δεδομένων, 2007).

2.5.2 Γνωσιακές προκαταλήψεις

Υπάρχει μια ποικιλία γνωστικών μεροληψιών που μπορούν να επηρεάσουν αρνητικά την ανάλυση. Για παράδειγμα, η προκατάληψη επιβεβαίωσης είναι η τάση αναζήτησης ή ερμηνείας πληροφοριών με τρόπο που να επιβεβαιώνει τις προκαταλήψεις του ατόμου. Επιπλέον, τα άτομα ενδέχεται να δυσφημήσουν τις πληροφορίες που δεν υποστηρίζουν τις απόψεις τους.

Οι αναλυτές μπορούν να εκπαιδευτούν ειδικά για να γνωρίζουν αυτές τις προκαταλήψεις και πώς να τις ξεπεράσουν. Επίσης, θα πρέπει να οριοθετούν σαφώς τις υποθέσεις τους και τις αλυσίδες συμπερασμάτων και να καθορίσουν το βαθμό και την πηγή

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας της αβεβαιότητας που εμπλέκονται στα συμπεράσματα (Παπαδημητρίου, Η Ανάλυση Δεδομένων-Παραγοντική Ανάλυση των Αντιστοιχιών - Ιεραρχική Ταξινόμηση και άλλες Μέθοδοι, 2007).

2.5.3 Αναξιπιστία

Οι αποτελεσματικοί αναλυτές είναι γενικά έμπειροι με μια ποικιλία αριθμητικών τεχνικών. Ωστόσο, όταν το κοινό μπορεί να μην έχει τέτοιου είδους γραφή με αριθμούς ή αριθμητική, λέγεται ότι είναι αμέτρητοι. Τα πρόσωπα που επικοινωνούν με τα δεδομένα ενδέχεται επίσης να προσπαθούν να παραπλανήσουν ή να παραπληροφήσουν, χρησιμοποιώντας σκόπιμα κακές αριθμητικές τεχνικές.

Για παράδειγμα, αν ένας αριθμός αυξάνεται ή πέφτει μπορεί να μην είναι ο βασικός παράγοντας. Σημαντικότερος μπορεί να είναι ο αριθμός σε σχέση με έναν άλλο αριθμό, όπως το μέγεθος των κρατικών εσόδων ή δαπανών σε σχέση με το μέγεθος της οικονομίας (ΑΕΠ) ή το ποσό του κόστους σε σχέση με τα έσοδα στις εταιρικές οικονομικές καταστάσεις. Αυτή η αριθμητική τεχνική αναφέρεται ως κανονικοποίηση ή κοινή μεγέθυνση. Υπάρχουν πολλές τέτοιες τεχνικές που χρησιμοποιούνται από τους αναλυτές, ανεξάρτητα από το εάν προσαρμόζονται για τον πληθωρισμό (π.χ. συγκρίνοντας τα πραγματικά με τα ονομαστικά δεδομένα) ή εξετάζοντας τις αυξήσεις του πληθυσμού, τα δημογραφικά στοιχεία κλπ. Οι αναλυτές εφαρμόζουν μια ποικιλία τεχνικών για την αντιμετώπιση των διαφόρων ποσοτικών μηνυμάτων που περιγράφονται στην παραπάνω ενότητα.

Οι αναλυτές μπορούν επίσης να αναλύσουν δεδομένα με διαφορετικές παραδοχές ή σενάρια. Για παράδειγμα, όταν οι αναλυτές πραγματοποιούν ανάλυση οικονομικών καταστάσεων, συχνά αναδιαμορφώνουν τις οικονομικές καταστάσεις με διαφορετικές παραδοχές για να βοηθήσουν στην εκτίμηση της μελλοντικής ταμειακής ροής, την οποία στη συνέχεια εκπίπτουν στην παρούσα αξία με βάση κάποιο επιτόκιο, για να καθορίσουν την αποτίμηση της εταιρείας ή το απόθεμά της (Παπαδημητρίου, Η Ανάλυση Δεδομένων-Παραγοντική Ανάλυση των Αντιστοιχιών - Ιεραρχική Ταξινόμηση και άλλες Μέθοδοι, 2007).

2.6 Ελεύθερα λογισμικά για την ανάλυση δεδομένων

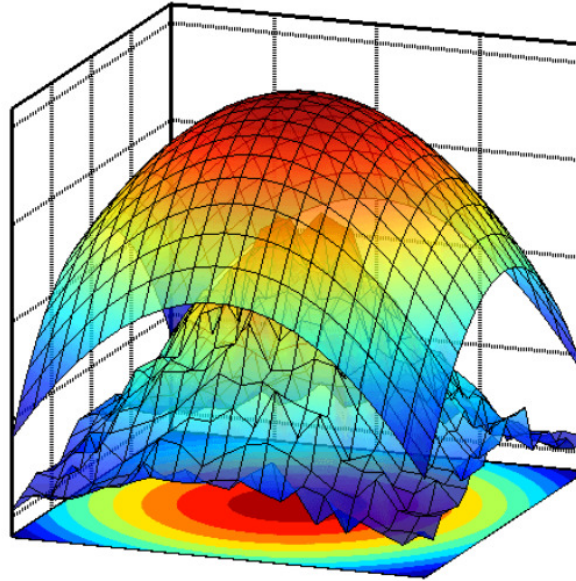
Στη συνέχεια, παρατίθενται διάφορα ελεύθερα λογισμικά για την ανάλυση δεδομένων:

1. DevInfo

Το DevInfo είναι ένα σύστημα βάσης δεδομένων που αναπτύχθηκε υπό την αιγίδα των Ηνωμένων Εθνών και εγκρίθηκε από την Ομάδα Ανάπτυξης των Ηνωμένων Εθνών για την παρακολούθηση της ανθρώπινης ανάπτυξης με συγκεκριμένο σκοπό την παρακολούθηση των Αναπτυξιακών Στόχων της Χιλιετίας (ΑΣΧ), το οποίο είναι ένα σύνολο από Δείκτες Ανθρώπινης Ανάπτυξης. Το DevInfo είναι ένα εργαλείο για την ομοιόμορφη οργάνωση, αποθήκευση και παρουσίαση δεδομένων για τη διευκόλυνση της ανταλλαγής δεδομένων σε εθνικό επίπεδο μεταξύ των κυβερνητικών υπηρεσιών, των υπηρεσιών του ΟΗΕ και των αναπτυξιακών εταιρών. Διανέμεται δωρεάν σε όλα τα κράτη μέλη του ΟΗΕ. Πρόκειται για περαιτέρω εξέλιξη του προηγούμενου συστήματος βάσης δεδομένων της UNICEF, ChildInfo (DevInfo, 2018).

2. DataMelt

Το DataMelt είναι ένα ελεύθερο λογισμικό για αριθμητικούς υπολογισμούς, μαθηματικά, στατιστικές, συμβολικούς υπολογισμούς, ανάλυση δεδομένων και οπτικοποίηση δεδομένων. Αυτό το multiplatform πρόγραμμα συνδυάζει την απλότητα των scripting languages, όπως Python, Ruby, Groovy, με τη δύναμη εκατοντάδων πακέτων Java (DataMelt Computation & Visualisation, 2018).

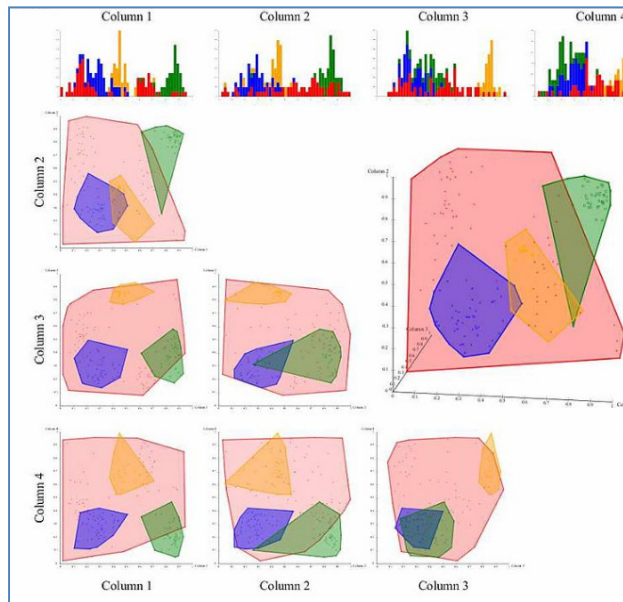


Σχήμα 2.5: Ανάλυση και οπτικοποίηση δεδομένων στο ελεύθερο λογισμικό για αριθμητικούς υπολογισμούς DataMelt.

Πηγή: (DataMelt Computation & Visualisation, 2018).

3. ELKI

Το ELKI (Environment for DeveLoping KDD-Applications Supported by Index-Structures) είναι μια ανακάλυψη γνώσεων σε βάσεις δεδομένων (KDD, «εξόρυξη δεδομένων»). Αυτό το λογισμικό αναπτύχθηκε για χρήση στην έρευνα και τη διδασκαλία, αρχικά στην ερευνητική μονάδα συστημάτων βάσεων δεδομένων του καθηγητή Hans-Peter Kriegel στο Πανεπιστήμιο Ludwig Maximilian του Μονάχου στη Γερμανία. Επιδιώκει να επιτρέψει την ανάπτυξη και αξιολόγηση προηγμένων αλγορίθμων εξόρυξης δεδομένων και την αλληλεπίδρασή τους με τις δομές δεικτών βάσεων δεδομένων (ELKI Data Mining, 2018).



Σχήμα 2.6: Ανάλυση και οπτικοποίηση δεδομένων στο ELKI.

Πηγή: (ELKI Data Mining, 2018).

4. KNIME

Το KNIME, το Konstanz Information Miner, είναι μια ελεύθερη και ανοιχτού κώδικα ανάλυση δεδομένων, πλατφόρμα αναφοράς και ενσωμάτωσης. Το KNIME ενσωματώνει διάφορα εξαρτήματα για την εκμάθηση μηχανών και την εξόρυξη δεδομένων μέσω της ιδέας της αγωγιμότητας των δεδομένων. Ένα γραφικό περιβάλλον χρήστη και η χρήση του JDBC επιτρέπει τη συναρμολόγηση κόμβων που συνδυάζουν διαφορετικές πηγές δεδομένων, συμπεριλαμβανομένης της προεπεξεργασίας (ETL - Extraction, Transformation, Loading), για μοντελοποίηση, ανάλυση δεδομένων και απεικόνιση χωρίς ή με ελάχιστο προγραμματισμό. Σε κάποιο βαθμό, το προηγμένο εργαλείο ανάλυσης KNIME μπορεί να θεωρηθεί ως εναλλακτική λύση (KNIME, 2018).

5. Orange

Η Orange είναι μια οπτικοποίηση δεδομένων ανοιχτού κώδικα, εργαλείο μάθησης μηχανών και εργαλείο εξόρυξης δεδομένων. Διαθέτει ένα front-end οπτικό προγραμματισμό για διερευνητική ανάλυση δεδομένων και οπτικοποίηση διαδραστικών δεδομένων και μπορεί επίσης να χρησιμοποιηθεί ως βιβλιοθήκη Python (Shaulsky, 2018).

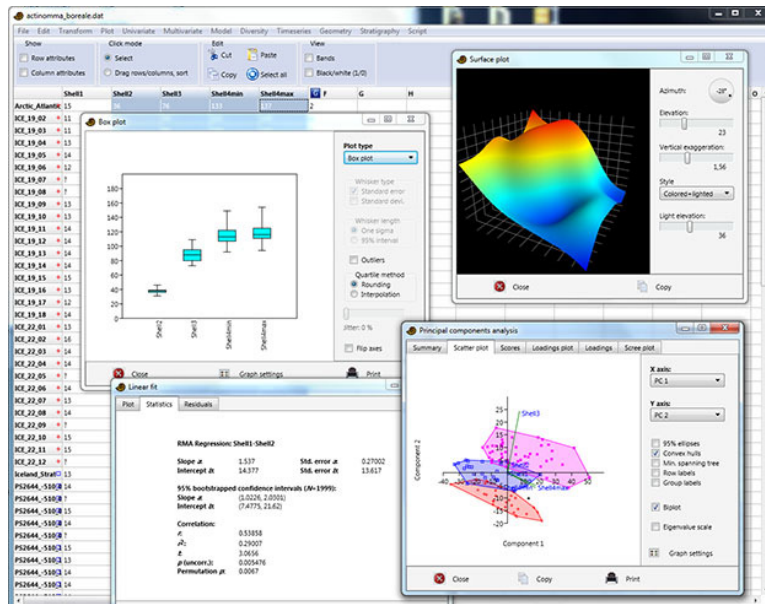


Σχήμα 2.7: Ανάλυση και οπτικοποίηση δεδομένων στην Orange.

Πηγή: (Shauly, 2018).

6. PAST

Το Past είναι ελεύθερο λογισμικό για την ανάλυση επιστημονικών δεδομένων, με λειτουργίες για χειρισμό δεδομένων, σχεδιασμό, μονομεταβλητές και πολυμεταβλητές στατιστικές, οικολογική ανάλυση, χρονοσειρές και χωρική ανάλυση, μορφομετρία και στρωματογραφία (Past of the Future, 2018).



Σχήμα 2.8: Ανάλυση και οπτικοποίηση δεδομένων στην Past.

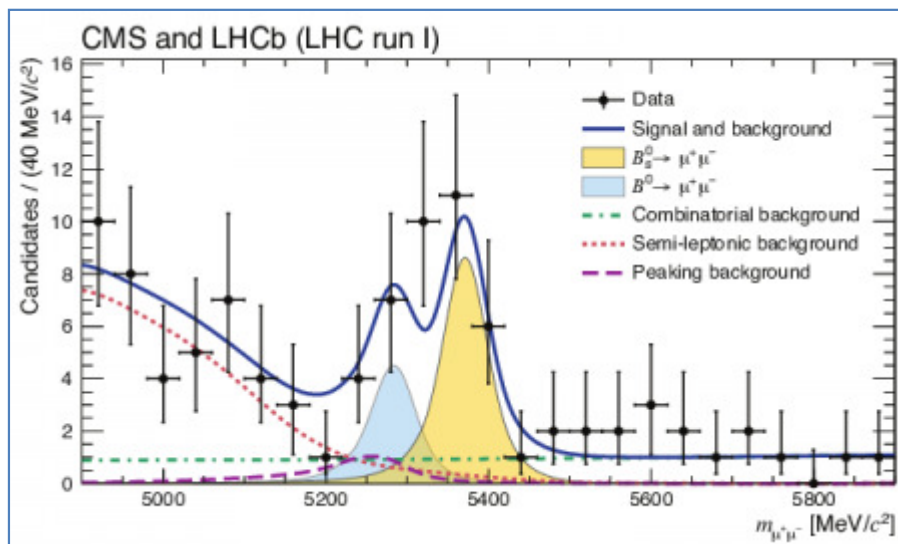
Πηγή: (Past of the Future, 2018).

7. «R»

«R» είναι γλώσσα προγραμματισμού και περιβάλλον ελεύθερου λογισμικού για στατιστικούς υπολογιστές και γραφικά που υποστηρίζονται από το Ίδρυμα Στατιστικής Πληροφορικής. Η γλώσσα «R» χρησιμοποιείται ευρέως μεταξύ στατιστικής και εξόρυξης δεδομένων για την ανάπτυξη στατιστικού λογισμικού αλλά και την ανάλυση δεδομένων (The R Foundation, 2018).

8. ROOT

Το ROOT είναι ένα αντικειμενοστρεφές πρόγραμμα και βιβλιοθήκη που αναπτύχθηκε από το CERN. Αρχικά σχεδιάστηκε για την ανάλυση δεδομένων σωματιδιακής φυσικής και περιέχει αρκετά χαρακτηριστικά ειδικά για αυτό το πεδίο, αλλά χρησιμοποιείται επίσης σε άλλες εφαρμογές όπως η αστρονομία και η εξόρυξη δεδομένων (CERN-ROOT Data Analysis Framework, 2018).

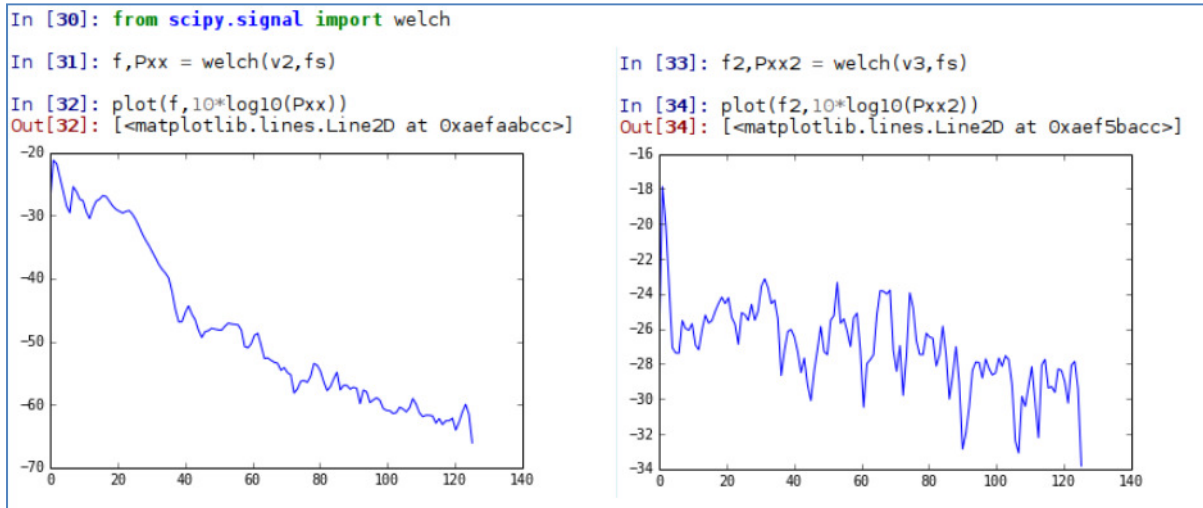


Σχήμα 2.9: Συνδυασμένα δεδομένα LHCb και CMS στο αντικειμενοστρεφές πρόγραμμα ROOT.

Πηγή: (CERN-ROOT Data Analysis Framework, 2018).

9. SciPy

Η SciPy είναι μια ελεύθερη -και ανοιχτού κώδικα- Python βιβλιοθήκη που χρησιμοποιείται σε επιστημονικούς υπολογιστές υπολογιστές. Περιέχει ενότητες βελτιστοποίησης, γραμμικής άλγεβρας, ολοκλήρωσης, ειδικές λειτουργίες, FFT, επεξεργασία σήματος και εικόνας (SciPy, 2019).

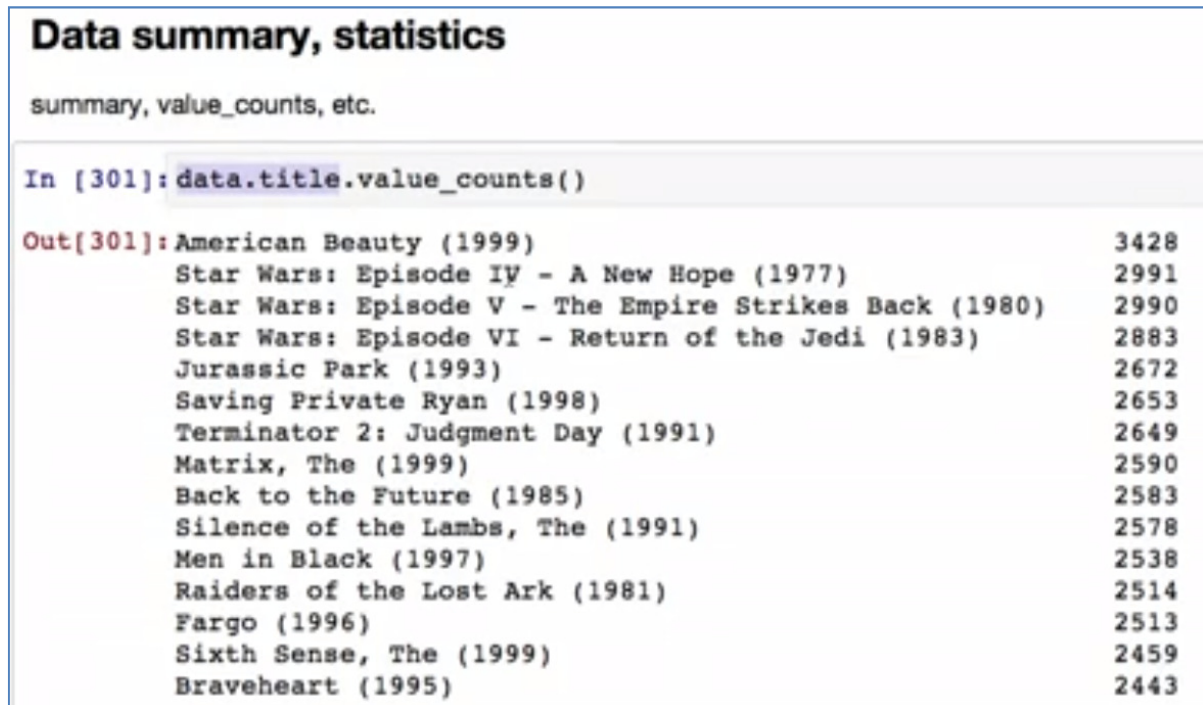


Σχήμα 2.10: Απεικόνιση δεδομένων PSD του ECG χρησιμοποιώντας scipy.

Πηγή: (SciPy, 2019).

10. Pandas

Η pandas είναι μια βιβλιοθήκη λογισμικού γραμμένη για τη γλώσσα προγραμματισμού Python για χειρισμό και ανάλυση δεδομένων. Συγκεκριμένα, προσφέρει δομές δεδομένων και λειτουργίες για τον χειρισμό αριθμητικών πινάκων και χρονοσειρών (pandas, 2018).



Σχήμα 2.11: Περιβάλλον ανάλυσης βιβλιοθήκη λογισμικού pandas.

Πηγή: (pandas, 2018).

3 Κεφάλαιο: «Οπτικοποίηση δεδομένων»

3.1 Γενικά στοιχεία

Η απεικόνιση δεδομένων θεωρείται από πολλούς κλάδους ως ένα σύγχρονο ισοδύναμο της οπτικής επικοινωνίας. Περιλαμβάνει τη δημιουργία και τη μελέτη της οπτικής αντιπροσώπευσης των δεδομένων.

Για να επικοινωνούν πληροφορίες με σαφήνεια και αποτελεσματικά, η οπτικοποίηση δεδομένων χρησιμοποιεί στατιστικά γραφήματα, διαγράμματα, γραφικά πληροφοριών και άλλα εργαλεία. Τα αριθμητικά δεδομένα μπορούν να κωδικοποιούνται χρησιμοποιώντας κουκκίδες, γραμμές ή ράβδους, για την οπτική επικοινωνία ενός ποσοτικού μηνύματος. Η αποτελεσματική απεικόνιση βοηθά τους χρήστες να αναλύουν και να αιτιολογούν τα δεδομένα και τα αποδεικτικά στοιχεία. Κάνει πολύπλοκα δεδομένα πιο προσιτά, κατανοητά και χρήσιμα. Οι χρήστες μπορούν να έχουν ιδιαίτερα αναλυτικά καθήκοντα, όπως είναι οι συγκρίσεις ή η κατανόηση της αιτιότητας, και η αρχή της σχεδίασης του γραφικού (δηλαδή παρουσιάζοντας συγκρίσεις ή δείχνοντας την αιτιότητα). Οι πίνακες χρησιμοποιούνται γενικά όπου οι χρήστες θα αναζητήσουν μια συγκεκριμένη μέτρηση, ενώ διαγράμματα διαφόρων τύπων χρησιμοποιούνται για την εμφάνιση μοτίβων ή σχέσεων στα δεδομένα για μία ή περισσότερες μεταβλητές(Μπακόλας & Τουρνόγλου, 2013).

Η απεικόνιση των δεδομένων είναι τέχνη και επιστήμη. Θεωρείται ως κλάδος περιγραφικών στατιστικών από ορισμένους, αλλά και ως εργαλείο ανάπτυξης θεμελιωμένων θεωριών από άλλους. Αυξημένα ποσά δεδομένων που δημιουργούνται από τη δραστηριότητα του Διαδικτύου και ένας αυξανόμενος αριθμός αισθητήρων στο περιβάλλον αναφέρονται ως «Big data / μεγάλα δεδομένα» ή «The Internet of Things (IoT) / Διαδίκτυο των πραγμάτων». Η επεξεργασία, η ανάλυση και η επικοινωνία αυτών των δεδομένων παρουσιάζουν δεοντολογικές και αναλυτικές προκλήσεις για την απεικόνιση των δεδομένων. Ο τομέας της επιστήμης των δεδομένων και των επαγγελματιών που καλούνται επιστήμονες δεδομένων βοηθά στην αντιμετώπιση αυτής της πρόκλησης(Ράτσης, 2009).

3.2 Ορισμοί / έννοιες

Η οπτικοποίηση δεδομένων αναφέρεται στις τεχνικές που χρησιμοποιούνται για την επικοινωνία δεδομένων ή πληροφοριών με την κωδικοποίησή τους ως οπτικά αντικείμενα

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας (για παράδειγμα, σημεία ή γραμμές) που περιέχονται σε γραφικά. Ο στόχος είναι η επικοινωνία των πληροφοριών με σαφήνεια και αποτελεσματικότητα στους χρήστες. Είναι ένα από τα βήματα στην ανάλυση δεδομένων ή την επιστήμη των δεδομένων. Σύμφωνα με τον Friedman (2008), ο «κύριος στόχος της απεικόνισης δεδομένων είναι η επικοινωνία των πληροφοριών με σαφήνεια και αποτελεσματικότητα μέσω γραφικών μέσων και δεν σημαίνει ότι η οπτικοποίηση δεδομένων πρέπει να φαίνεται βαρετή να είναι λειτουργική ή εξαιρετικά περίπλοκη για να φανεί όμορφη, παρέχοντας πληροφορίες για ένα αρκετά αραιό και περίπλοκο σύνολο δεδομένων, γνωρίζοντας με τον πιο διαισθητικό τρόπο τις βασικές πτυχές της. Ωστόσο, συχνά οι σχεδιαστές δεν επιτυγχάνουν ισορροπία μεταξύ μορφής και λειτουργίας, δημιουργώντας οπτικοποιήσεις δεδομένων που δεν εξυπηρετούν τον κύριο σκοπό τους - να κοινοποιούν πληροφορίες» (Hoelscher, 2018).

Πράγματι, οι Fernanda Viegas και Martin M. Wattenberg πρότειναν ότι η ιδανική απεικόνιση δεν πρέπει μόνο να επικοινωνεί με σαφήνεια, αλλά να διεγείρει τη συμμετοχή και την προσοχή των θεατών.

Η οπτικοποίηση δεδομένων είναι στενά συνδεδεμένη με τα γραφικά πληροφοριών, την οπτικοποίηση πληροφοριών, την επιστημονική απεικόνιση, τη διερευνητική ανάλυση δεδομένων και τα στατιστικά γραφικά. Στη νέα χιλιετία, η απεικόνιση δεδομένων έχει καταστεί ενεργός τομέας έρευνας, διδασκαλίας και ανάπτυξης (Sewell, 2018).

3.3 Χαρακτηριστικά των αποτελεσματικών γραφικών απεικονίσεων

Ο καθηγητής Edward Tufte (πρωτοπόρος στον τομέα της απεικόνισης δεδομένων) εξήγησε ότι οι χρήστες των γραφικών απεικονίσεων εκτελούν συγκεκριμένα αναλυτικά καθήκοντα όπως η πραγματοποίηση συγκρίσεων. Η αρχή σχεδιασμού του γραφικού πληροφοριών πρέπει να υποστηρίζει την αναλυτική εργασία. Όπως δείχνουν οι William Cleveland και Robert McGill, διάφορα γραφικά στοιχεία το επιτελούν περισσότερο ή λιγότερο αποτελεσματικά (Hoelscher, 2018).

Στο βιβλίο του 1983 «The Visual Display of Quantitative Information / Η οπτική απεικόνιση των ποσοτικών πληροφοριών», ο Edward Tufte ορίζει τις γραφικές απεικονίσεις και τις αρχές για την αποτελεσματική γραφική απεικόνιση στο ακόλουθο απόσπασμα (Ξυνταράκης & Αντωνίου, 2019)

«Η αριστεία στα στατιστικά γραφικά αποτελείται από σύνθετες ιδέες που επικοινωνούν με σαφήνεια, ακρίβεια και αποδοτικότητα». Οι αρχές τους είναι οι εξής:

- Η εμφάνιση των δεδομένων
- Η παρακίνηση του θεατή να σκεφτεί για την ουσία και όχι για τη μεθοδολογία, τον γραφικό σχεδιασμό, την τεχνολογία της γραφικής παραγωγής ή κάτι άλλο
- Η αποφυγή της στρέβλωσης του τι πρέπει να πουν τα δεδομένα
- Η παρουσίαση πολλών αριθμών σε ένα μικρό χώρο
- Η συνοχή μεταξύ των μεγάλων συνόλων δεδομένων
- Η ενθάρρυνση όσον αφορά τη σύγκριση διαφορετικών στοιχείων
- Η εμφάνιση των δεδομένων σε διάφορα επίπεδα λεπτομέρειας, από μια ευρεία επισκόπηση έως την ωραία δομή
- Η εξυπηρέτηση ενός εύλογα σαφούς σκοπού, όπως περιγραφή, εξερεύνηση, χαρτογράφηση ή διακόσμηση
- Η ενσωμάτωση στις στατιστικές και λεκτικές περιγραφές ενός συνόλου δεδομένων.

Τα γραφικά αποκαλύπτουν δεδομένα. Πράγματι, τα γραφικά μπορούν να είναι πιο ακριβή και αποκαλυπτικά από τους συμβατικούς στατιστικούς υπολογισμούς».

Η μη εφαρμογή αυτών των αρχών μπορεί να οδηγήσει σε παραπλανητικά γραφήματα, τα οποία διαστρεβλώνουν το μήνυμα ή στηρίζουν ένα εσφαλμένο συμπέρασμα. Σύμφωνα με τον Tufte, το chartjunk αναφέρεται στην ξένη εσωτερική διακόσμηση του γραφικού που δεν ενισχύει το μήνυμα, ή στις άχρηστες τρισδιάστατες επιπτώσεις. Διαχωρίζοντας άσκοπα το επεξηγηματικό κλειδί από την ίδια την εικόνα, απαιτώντας από το μάτι να ταξιδεύει εμπρός και πίσω από την εικόνα στο κλειδί, είναι μια μορφή «διοικητικών συντριμμάτων (Ξυνταράκης & Αντωνίου, 2019).

3.4 Ποσοτικά μηνύματα

Σύμφωνα με την διεθνή βιβλιογραφία περιγράφονται οκτώ τύποι ποσοτικών μηνυμάτων που οι χρήστες ενδέχεται να επιχειρήσουν να κατανοήσουν ή να επικοινωνήσουν από ένα σύνολο δεδομένων. Τα σχετικά γραφήματα που χρησιμοποιούνται για να βοηθήσουν στην επικοινωνία του μηνύματος είναι (Sewell, 2018):

1. **Χρονολογικές σειρές:** Μια ενιαία μεταβλητή καταγράφεται σε μια χρονική περίοδο, όπως το ποσοστό ανεργίας σε μια περίοδο 10 ετών. Μπορεί να χρησιμοποιηθεί ένα γράφημα γραμμής για να αποδειχθεί η τάση.
2. **Κατάταξη:** Οι κατηγοριοποιημένες υποδιαιρέσεις κατατάσσονται σε αύξουσα ή φθίνουσα σειρά, όπως η κατάταξη των επιδόσεων των πωλήσεων (το μέτρο) από τους

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

πωλητές (την κατηγορία, με κάθε πωλητή κατηγοριοποιημένη υποδιαίρεση) κατά τη διάρκεια μιας μόνο περιόδου. Ένα διάγραμμα ράβδων μπορεί να χρησιμοποιηθεί για να δείξει τη σύγκριση μεταξύ των ατόμων πωλήσεων.

3. **Μέρος σε ολόκληρο:** Οι κατηγοριοποιημένες υποδιαίρεσεις μετρούνται σε σχέση με το σύνολο (δηλαδή ένα ποσοστό από το 100%). Ένα διάγραμμα πίτας ή ένα διάγραμμα ράβδων μπορεί να δείξει τη σύγκριση των λόγων, όπως το μερίδιο αγοράς που αντιπροσωπεύουν οι ανταγωνιστές σε μια αγορά.
4. **Απόκλιση:** Οι κατηγοριοποιημένες υποδιαίρεσεις συγκρίνονται με μια αναφορά, όπως μια σύγκριση των πραγματικών έναντι των δαπανών του προϋπολογισμού για διάφορα τμήματα μιας επιχείρησης για μια δεδομένη χρονική περίοδο. Ένα διάγραμμα ράβδων μπορεί να δείξει σύγκριση του πραγματικού με το ποσό αναφοράς.
5. **Κατανομή συχνότητας:** Δείχνει τον αριθμό των παρατηρήσεων μιας συγκεκριμένης μεταβλητής για δεδομένο διάστημα, όπως ο αριθμός των ετών κατά τα οποία η χρηματιστηριακή απόδοση είναι μεταξύ διαστημάτων όπως 0-10%, 11-20% κλπ. Ένα ιστόγραμμα, ένας τύπος διαγράμματος ράβδων, μπορεί να χρησιμοποιηθεί για την ανάλυση αυτή. Μια boxplot βοηθά στην απεικόνιση βασικών στατιστικών στοιχείων σχετικά με τη διανομή, όπως διάμεσος, τεταρτημόρια, ακραίες τιμές, κ.λπ.
6. **Συσχέτιση:** Σύγκριση μεταξύ παρατηρήσεων που παριστάνονται από δύο μεταβλητές (X, Y) για να προσδιοριστεί εάν τείνουν να κινούνται στις ίδιες ή αντίθετες κατευθύνσεις. Για παράδειγμα, σχεδιάζεται η ανεργία (X) και ο πληθωρισμός (Y) για ένα δείγμα μηνών. Για αυτό το μήνυμα χρησιμοποιείται συνήθως μια γραφική παράσταση σκέδασης.
7. **Ονομαστική σύγκριση:** Συγκρίνοντας κατηγορηματικές υποδιαίρεσεις σε καμία συγκεκριμένη σειρά, όπως ο όγκος των πωλήσεων ανά κωδικό προϊόντος. Για τη σύγκριση αυτή μπορεί να χρησιμοποιηθεί ένα διάγραμμα ράβδων.
8. **Γεωγραφική ή γεωχωρική:** Σύγκριση μιας μεταβλητής σε έναν χάρτη, όπως το ποσοστό ανεργίας ανά κράτος ή ο αριθμός των ατόμων στους διάφορους ορόφους ενός κτιρίου. Ένα χαρτογράφημα είναι ένα τυπικό γραφικό που χρησιμοποιείται.

Οι αναλυτές που εξετάζουν ένα σύνολο δεδομένων ενδέχεται να εξετάσουν εάν ορισμένα ή όλα τα παραπάνω μηνύματα και τύποι γραφικών ισχύουν για το έργο και το κοινό τους. Η διαδικασία δοκιμής και σφάλματος για τον προσδιορισμό σημαντικών σχέσεων και μηνυμάτων στα δεδομένα αποτελεί μέρος της διερευνητικής ανάλυσης δεδομένων.

3.5 Οπτική αντίληψη και απεικόνιση δεδομένων

Ένας άνθρωπος μπορεί να διακρίνει εύκολα τις διαφορές στο μήκος, το σχήμα, τον προσανατολισμό και το χρώμα (απόχρωση) χωρίς σημαντική προσπάθεια επεξεργασίας. Αυτά αναφέρονται ως «προσεκτικά χαρακτηριστικά» (η προσεκτική επεξεργασία είναι η υποσυνείδητη συσσώρευση πληροφοριών από το περιβάλλον και όλες οι διαθέσιμες πληροφορίες υποβάλλονται σε προσεκτική επεξεργασία). Για παράδειγμα, μπορεί να απαιτηθεί σημαντικός χρόνος και προσπάθεια («προσεκτική επεξεργασία») για τον προσδιορισμό του αριθμού των ψηφίων «5» που εμφανίζονται σε μια σειρά αριθμών, αλλά εάν το ψηφίο αυτό έχει διαφορετικό μέγεθος, προσανατολισμό ή χρώμα, οι στιγμές του ψηφίου μπορούν να σημειωθούν γρήγορα μέσω της προληπτικής επεξεργασίας (Ράτσης, 2009).

Τα αποτελεσματικά γραφικά εκμεταλλεύονται την προ-προσεκτική επεξεργασία και τα χαρακτηριστικά και τη σχετική ισχύ αυτών των χαρακτηριστικών. Για παράδειγμα, επειδή ο άνθρωπος μπορεί να επεξεργάζεται ευκολότερα τις διαφορές στο μήκος της γραμμής από την επιφάνεια, μπορεί να είναι πιο αποτελεσματικό να χρησιμοποιείται ένα διάγραμμα ράβδων (το οποίο εκμεταλλεύεται το μήκος γραμμής για να δείξει σύγκριση) αντί για διαγράμματα πίτας.

Σχεδόν όλες οι οπτικοποιήσεις δεδομένων δημιουργούνται για ανθρώπινη κατανάλωση. Η γνώση της ανθρώπινης αντίληψης και της γνώσης είναι απαραίτητη όταν σχεδιάζονται διαισθητικές απεικονίσεις. Η ανθρώπινη οπτική επεξεργασία είναι αποτελεσματική στην αντίληψη αλλαγών και στην πραγματοποίηση συγκρίσεων μεταξύ των ποσοτήτων, των μεγεθών, των σχημάτων και των παραλλαγών της ελαφρότητας. Όταν οι ιδιότητες των συμβολικών δεδομένων χαρτογραφούνται στις οπτικές ιδιότητες, ο άνθρωπος μπορεί να αναζητήσει αποτελεσματικά μεγάλες ποσότητες δεδομένων. Εκτιμάται ότι τα 2/3 των νευρώνων του εγκεφάλου μπορούν να συμμετέχουν στην οπτική επεξεργασία. Η σωστή απεικόνιση παρέχει μια διαφορετική προσέγγιση για να δείξει πιθανές συνδέσεις, σχέσεις κ.λπ. οι οποίες δεν είναι τόσο προφανείς σε μη οπτικοποιημένα ποσοτικά δεδομένα. Ως εκ τούτου η απεικόνιση μπορεί να γίνει μέσο έρευνας (Παπαδημητρίου, Η ανάλυση δεδομένων, 2007).

3.6 Ιστορική αναδρομή της απεικόνισης δεδομένων

Η πρώτη τεκμηριωμένη απεικόνιση δεδομένων μπορεί να εντοπιστεί πίσω στο 1160 π.Χ. με το χάρτη Papyrus του Τορίνο που απεικονίζει με ακρίβεια την κατανομή των γεωλογικών πόρων και παρέχει πληροφορίες για την εξόρυξη αυτών των πόρων. Αυτοί οι χάρτες μπορούν να κατηγοριοποιηθούν ως θεματική χαρτογραφία, η οποία είναι ένας τύπος απεικόνισης δεδομένων που παρουσιάζει και κοινοποιεί συγκεκριμένα δεδομένα και πληροφορίες μέσω μιας γεωγραφικής απεικόνισης που αποσκοπεί στην εμφάνιση ενός συγκεκριμένου θέματος που συνδέεται με μια συγκεκριμένη γεωγραφική περιοχή. Οι πρώτες τεκμηριωμένες μορφές απεικόνισης δεδομένων ήταν διάφοροι θεματικοί χάρτες από διαφορετικούς πολιτισμούς και ιδεογράμματα και ιερογλυφικά που παρείχαν και επέτρεψαν την ερμηνεία των πληροφοριών που απεικονίζονται. Για παράδειγμα, τα δισκία Γραμμικής Β' των Μυκηνών παρέσχε μια απεικόνιση των πληροφοριών σχετικά με την εποχή της Ύστερης Εποχής του Χαλκού στη Μεσόγειο. Η ιδέα των συντεταγμένων χρησιμοποιήθηκε από τους αρχαίους Αιγύπτιους επιθεωρητές για την τοποθέτηση πόλεων, οι γήινες και ουράνιες θέσεις βρισκόταν από κάτι παρόμοιο με το γεωγραφικό πλάτος και το γεωγραφικό μήκος τουλάχιστον κατά το 200 π.Χ. και η προβολή του χάρτη μιας σφαιρικής γης σε γεωγραφικό πλάτος και γεωγραφικό μήκος από τον Κλαύδιο Πτολεμαίο (85 – 165 π.Χ.) στην Αλεξάνδρεια θα χρησίμευε ως πρότυπα αναφοράς μέχρι τον 14^ο αιώνα (Χαριδάκου - Κεραμίδα, 2013).

Η εφεύρεση χαρτιού και περγαμινών επέτρεψε την περαιτέρω ανάπτυξη απεικονίσεων σε όλη την ιστορία.

Μέχρι τον 16^ο αιώνα, αναπτύχθηκαν καλά τεχνικές και όργανα για ακριβή παρατήρηση και μέτρηση φυσικών μεγεθών, για γεωγραφική και ουράνια θέση όπως για παράδειγμα ένα «τεταρτημόριο τοίχου» που κατασκευάστηκε από τον Tycho Brahe (1546-1601) στο παρατηρητήριο του. Ιδιαίτερα σημαντική ήταν η ανάπτυξη τριγωνισμού και άλλων μεθόδων για τον ακριβή προσδιορισμό των θέσεων χαρτογράφησης (Χαριδάκου - Κεραμίδα, 2013).

Στο δεύτερο μισό του 20ού αιώνα, ο Jacques Bertin χρησιμοποίησε ποσοτικά γραφήματα για να αναπαριστά πληροφορίες «διαισθητικά, καθαρά, με ακρίβεια και αποτελεσματικά».

Ο John Tukey και ο Edward Tufte ανέδειξαν τα όρια της απεικόνισης δεδομένων. Ο Tukey με τη νέα στατιστική του προσέγγιση για διερευνητική ανάλυση δεδομένων και ο

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

Tufte με το βιβλίο του «*The visual representation of quantitative information / Η οπτική απεικόνιση των ποσοτικών πληροφοριών*» άνοιξε το δρόμο για τη βελτίωση των τεχνικών οπτικοποίησης δεδομένων για περισσότερους από τους στατιστικούς. Με την εξέλιξη της τεχνολογίας ήρθε η εξέλιξη της οπτικοποίησης δεδομένων, ξεκινώντας απεικονίσεις με το χέρι και εξελισσόμενες σε πιο τεχνικές εφαρμογές - συμπεριλαμβανομένων διαδραστικών σχεδίων που οδηγούν στην απεικόνιση του λογισμικού (Βλαχάβας, Κεφαλάς, Βασιλειάδης, Κόκκορας, & Σακελλαρίου, 2011).

Προγράμματα όπως το SAS, SOFA, R, Minitab, Cornerstone επιτρέπουν την απεικόνιση δεδομένων στον τομέα των στατιστικών. Άλλες εφαρμογές οπτικοποίησης δεδομένων, πιο εστιασμένες και μοναδικές για τα άτομα, οι γλώσσες προγραμματισμού όπως το D3, το Python και το JavaScript συμβάλλουν στη δυνατότητα απεικόνισης των ποσοτικών δεδομένων. Ιδιωτικά σχολεία έχουν επίσης αναπτύξει προγράμματα για να καλύψουν τη ζήτηση για οπτικοποίηση μαθησιακών δεδομένων και τις σχετικές βιβλιοθήκες προγραμματισμού, συμπεριλαμβανομένων δωρεάν προγραμμάτων όπως το The Incubator ή προπληρωμένα προγράμματα όπως η General Assembly (Sewell, 2018).

3.7 Έννοιες – ορισμοί

Η απεικόνιση δεδομένων περιλαμβάνει συγκεκριμένη ορολογία και έννοιες από τις οποίες μερικές προέρχονται από στατιστικές. Για παράδειγμα, ο συντάκτης Stephen Few ορίζει δύο τύπους δεδομένων, τα οποία χρησιμοποιούνται σε συνδυασμό για να υποστηρίξουν μια ουσιαστική ανάλυση ή οπτικοποίηση (Χαριδάκου - Κεραμίδα, 2013):

1. «**Κατηγορία**»: Ετικέτες κειμένου που περιγράφουν τη φύση των δεδομένων, όπως «Όνομα» ή «Ηλικία». Ο όρος αυτός καλύπτει επίσης ποιοτικά (μη αριθμητικά) δεδομένα.
2. «**Ποσοτικά**»: Αριθμητικά μέτρα, όπως «30» για να αντιπροσωπεύουν την ηλικία σε χρόνια.

Δύο κύριοι τύποι προβολών πληροφοριών είναι πίνακες και γραφήματα (Χαριδάκου - Κεραμίδα, 2013).

- a) Ένας **πίνακας** περιέχει ποσοτικά δεδομένα οργανωμένα σε σειρές και στήλες με κατηγοριοποιημένες ετικέτες. Χρησιμοποιείται κυρίως για την αναζήτηση συγκεκριμένων τιμών. Για παράδειγμα, ένας πίνακας μπορεί να έχει κατηγοριοποιημένες ετικέτες στηλών που αντιπροσωπεύουν το όνομα (ποιοτική μεταβλητή) και την ηλικία (ποσοτική

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

μεταβλητή), με κάθε σειρά δεδομένων να αντιπροσωπεύει ένα άτομο (την πειραματική μονάδα ή την υποδιαίρεση κατηγορίας).

- b) Ένα **γράφημα** χρησιμοποιείται κυρίως για την εμφάνιση σχέσεων μεταξύ των δεδομένων και των εικόνων που κωδικοποιούνται ως οπτικά αντικείμενα (π.χ. γραμμές ή σημεία). Οι αριθμητικές τιμές εμφανίζονται σε μια περιοχή που οριοθετείται από έναν ή περισσότερους άξονες. Αυτοί οι άξονες παρέχουν κλίμακες (ποσοτικές και κατηγορικές) που χρησιμοποιούνται για την επισήμανση και την εκχώρηση αξιών στα οπτικά αντικείμενα. Πολλά γραφήματα αναφέρονται επίσης ως διαγράμματα.

Οι Erppler και Lengler έχουν αναπτύξει τον «*Περιοδικό πίνακα μεθόδων απεικόνισης*», ένα διαδραστικό χάρτη που παρουσιάζει διάφορες μεθόδους οπτικοποίησης δεδομένων. Περιλαμβάνει έξι τύπους μεθόδων οπτικοποίησης δεδομένων: δεδομένα, πληροφορίες, έννοια, στρατηγική, μεταφορά και ένωση.

3.8 Παραδείγματα διαγραμμάτων χρήσης οπτικοποίησης δεδομένων

Παραδείγματα διαγραμμάτων χρήσης οπτικοποίησης δεδομένων καταγράφονται παρακάτω:

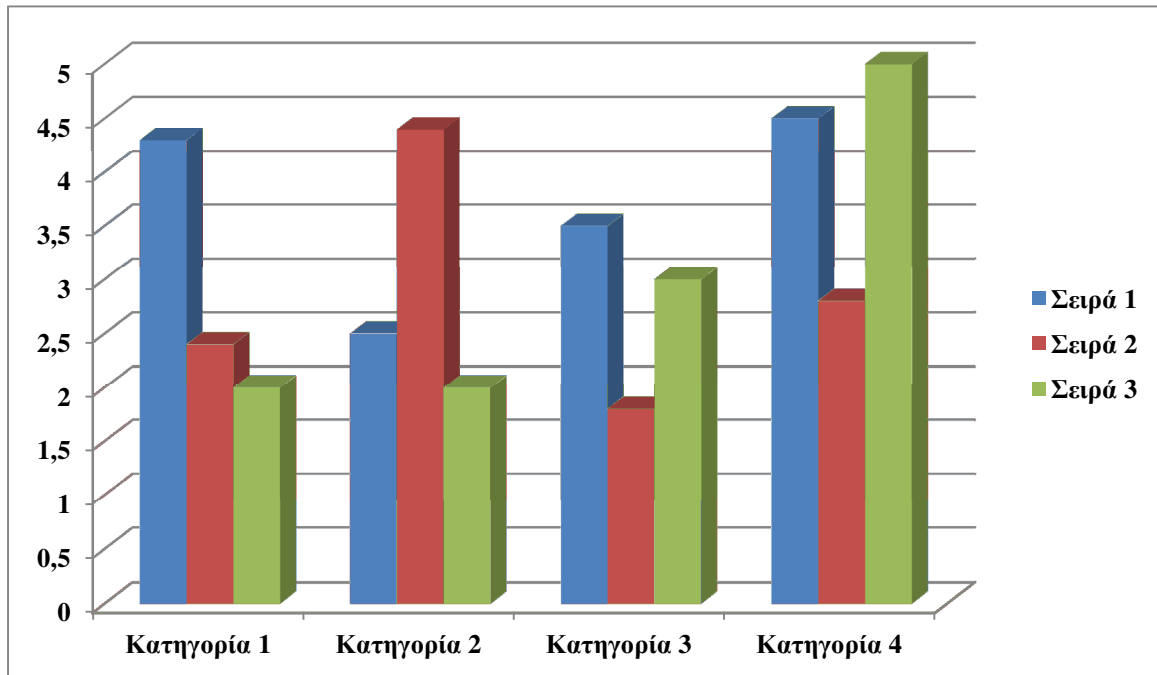
3.8.1 Ραβδόγραμμα

Ένα διάγραμμα ράβδων ή ένα γράφημα γραμμών είναι ένα διάγραμμα ή ένα γράφημα που παρουσιάζει τα κατηγορηματικά δεδομένα με ορθογώνιες ράβδους με ύψη ή μήκη ανάλογα με τις τιμές που αντιπροσωπεύουν. Οι ράβδοι μπορούν να γραφιστούν κατακόρυφα ή οριζόντια. Ένα γράφημα κάθετης ράβδου καλείται μερικές φορές γραμμικό γράφημα.

Ένα διάγραμμα ράβδων παρουσιάζει συγκρίσεις μεταξύ διακριτών κατηγοριών. Ένας άξονας του διαγράμματος δείχνει τις συγκεκριμένες κατηγορίες που συγκρίνονται και ο άλλος άξονας αντιπροσωπεύει μια μετρούμενη τιμή. Ορισμένες γραμμικές γραφικές παραστάσεις παρουσιάζουν ράβδους συγκεντρωμένες σε ομάδες περισσότερων της μιας, που δείχνουν τις τιμές περισσότερων από μία μετρημένων μεταβλητών (Γιάκας, 2013).

Οι διαγράμματα γραμμών έχουν διακριτή περιοχή κατηγοριών και συνήθως κλιμακώνονται έτσι ώστε όλα τα δεδομένα να μπορούν να χωρέσουν στο γράφημα. Όταν δεν υπάρχει φυσική ταξινόμηση των κατηγοριών που συγκρίνονται, οι ράβδοι στο διάγραμμα μπορούν να ταξινομηθούν με οποιαδήποτε σειρά. Διαγράμματα ράβδων διατεταγμένα από την υψηλότερη έως τη χαμηλότερη συχνότητα ονομάζονται διαγράμματα Pareto. Τα γραφήματα / διαγράμματα παρέχουν μια οπτική παρουσίαση κατηγορικών δεδομένων.

Κατηγορικά δεδομένα είναι μια ομαδοποίηση δεδομένων σε διακριτές ομάδες, όπως μήνες του έτους, ηλικιακή ομάδα, μεγέθη παπουτσιών και ζώα. Αυτές οι κατηγορίες είναι συνήθως ποιοτικές. Σε ένα γράφημα ράβδων στήλης, οι κατηγορίες εμφανίζονται κατά μήκος του οριζώντιου άξονα και το ύψος της ράβδου αντιστοιχεί στην αξία κάθε κατηγορίας (Σιάρδος, 2017).



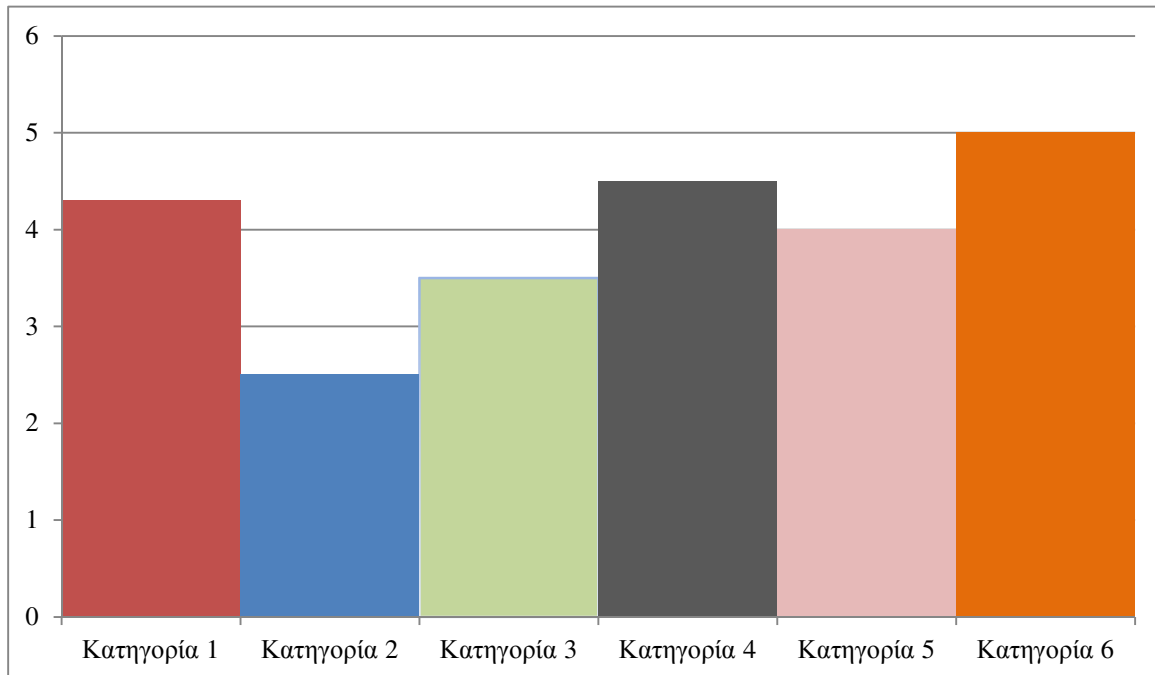
Σχήμα 3.1: Παράδειγμα διαγράμματος ράβδων.

3.8.2 Ιστόγραμμα

Ένα ιστόγραμμα είναι μια ακριβής αναπαράσταση της κατανομής αριθμητικών δεδομένων. Πρόκειται για μια εκτίμηση της κατανομής πιθανότητας μίας συνεχούς μεταβλητής (ποσοτική μεταβλητή) και εισήχθη για πρώτη φορά από τον Karl Pearson. Διαφέρει από ένα γράφημα ράβδων, με την έννοια ότι ένα γράφημα γραμμών αντιστοιχεί σε δύο μεταβλητές, αλλά ένα ιστόγραμμα αφορά μόνο ένα (Σιάρδος, 2017). Για την κατασκευή ενός ιστογράμματος, το πρώτο βήμα είναι η ομαδοποίηση δεδομένων / το εύρος τιμών - δηλαδή, η διαίρεση ολόκληρου του φάσματος τιμών σε μια σειρά διαστημάτων - και στη συνέχεια η μέτρηση των τιμών που πέφτουν σε κάθε διάστημα. Οι κάδοι συνήθως ορίζονται ως διαδοχικά, μη επικαλυπτόμενα διαστήματα μιας μεταβλητής. Οι κάδοι (διαστήματα) πρέπει να είναι διπλά και συχνά (αλλά δεν απαιτείται να είναι) ίσου μεγέθους (Γιάκας, 2013).

Εάν οι κάδοι είναι ίσου μεγέθους, ένα ορθογώνιο τοποθετείται πάνω από τον κάδο με ύψος ανάλογο με τη συχνότητα - τον αριθμό των περιπτώσεων σε κάθε κάδο. Ένα ιστόγραμμα μπορεί επίσης να ομαλοποιηθεί για να εμφανίσει τις «σχετικές» συχνότητες. Στη

συνέχεια, δείχνει το ποσοστό των περιπτώσεων που εμπίπτουν σε καθεμία από τις διάφορες κατηγορίες, με το άθροισμα των υψών να ισούται με 1.



Σχήμα 3.2: Παράδειγμα «Πολυτροπικού» ιστογράμματος.

Ωστόσο, οι κάδοι δεν πρέπει να έχουν ίσο πλάτος. Στην περίπτωση αυτή, το ορθογώνιο που έχει ανεγερθεί ορίζεται ότι έχει την περιοχή του ανάλογη με τη συχνότητα των περιπτώσεων στον κάδο. Ο κάθετος άξονας δεν είναι τότε η συχνότητα αλλά η πυκνότητα συχνότητας - ο αριθμός των περιπτώσεων ανά μονάδα της μεταβλητής στον οριζόντιο άξονα.

Καθώς οι γειτονικοί κάδοι δεν αφήνουν κενά, τα ορθογώνια ενός ιστογράμματος αγγίζουν το ένα το άλλο για να δείξουν ότι η αρχική μεταβλητή είναι συνεχής.

Τα ιστογράμματα δίνουν μια αίσθηση της πυκνότητας της υποκείμενης κατανομής των δεδομένων και συχνά για την εκτίμηση της πυκνότητας. Η συνολική περιοχή ενός ιστογράμματος που χρησιμοποιείται για την πυκνότητα πιθανότητας κανονικοποιείται πάντοτε στο 1. Εάν το μήκος των διαστημάτων στον x- άξονα είναι 1, τότε ένα ιστόγραμμα είναι ταυτόσημο με ένα σχετικό διάγραμμα συχνότητας (Σιάρδος, 2017).

Ένα ιστόγραμμα μπορεί να θεωρηθεί ως μια απλοϊκή εκτίμηση πυκνότητας πυρήνα, η οποία χρησιμοποιεί έναν πυρήνα για να εξομαλύνει τις συχνότητες πάνω από τους κάδους. Αυτό αποδίδει μια πιο ομαλή συνάρτηση πυκνότητας πιθανότητας, η οποία σε γενικές γραμμές θα αντικατοπτρίζει καλύτερα την κατανομή της υποκείμενης μεταβλητής. Η εκτίμηση της πυκνότητας μπορεί να γραφεί ως εναλλακτική λύση στο ιστόγραμμα και

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

συνήθως συνιστάται ως καμπύλη παρά ως σύνολο κιβωτίων. Ωστόσο, τα ιστογράμματα προτιμώνται σε εφαρμογές, όταν οι στατιστικές ιδιότητες τους πρέπει να διαμορφώνονται. Η συσχτισμένη παραλλαγή της εκτίμησης πυκνότητας πυρήνα είναι πολύ δύσκολο να περιγραφεί μαθηματικά, ενώ είναι απλή για ένα ιστόγραμμα όπου κάθε κάδος ποικίλει ανεξάρτητα.

Μια εναλλακτική λύση για την εκτίμηση της πυκνότητας του πυρήνα είναι το μέσο ιστόγραμμα μετατόπισης το οποίο είναι γρήγορο για υπολογισμό και δίνει μια εκτίμηση της ομαλής καμπύλης της πυκνότητας χωρίς τη χρήση πυρήνων.

Το ιστόγραμμα είναι ένα από τα επτά βασικά εργαλεία ελέγχου ποιότητας³.

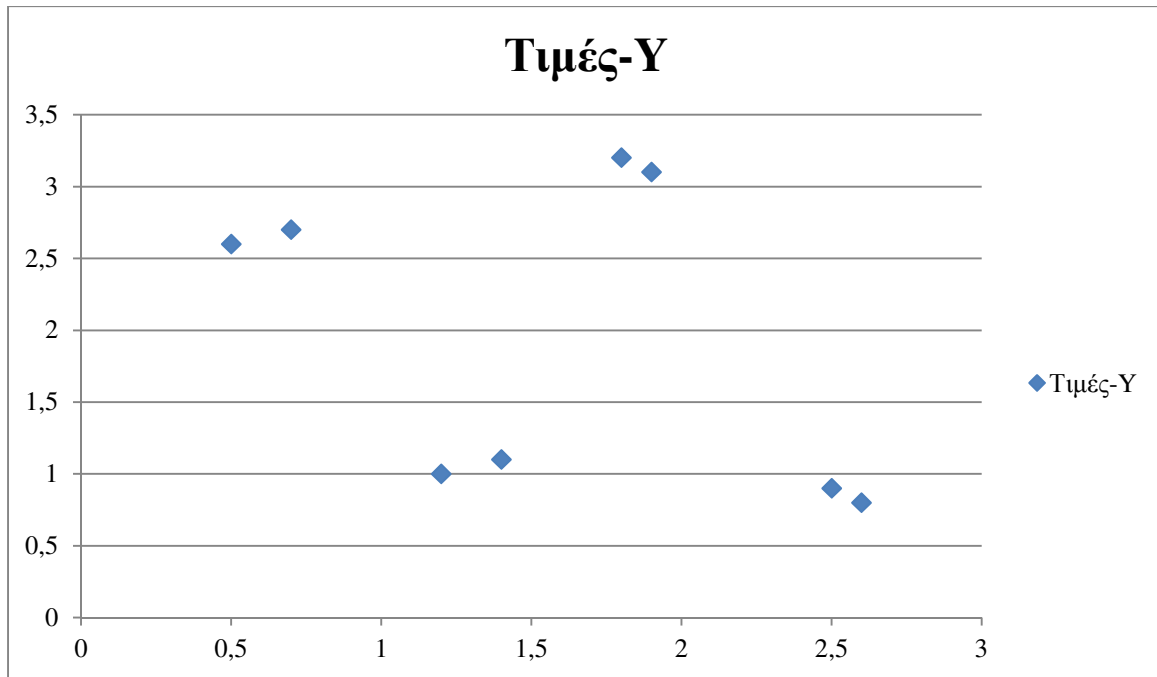
Τα ιστογράμματα μερικές φορές συγχέονται με τα γραφήματα. Ένα ιστόγραμμα χρησιμοποιείται για συνεχή δεδομένα, όπου οι κάδοι αντιπροσωπεύουν εύρη δεδομένων, ενώ ένα διάγραμμα ράβδων είναι μια γραφική παράσταση κατηγορικών μεταβλητών. Ορισμένοι συγγραφείς συνιστούν ότι τα διαγράμματα bar έχουν κενά μεταξύ των ορθογωνίων για να διευκρινιστεί η διάκριση (Σιάρδος, 2017).

3.8.3 Διάγραμμα διασποράς

Ένα διάγραμμα σκέδασης / διασποράς (scatterplot) είναι ένας τύπος γραφήματος ή μαθηματικού διαγράμματος που χρησιμοποιεί καρτεσιανές συντεταγμένες για την εμφάνιση τιμών για τυπικά δύο μεταβλητές για ένα σύνολο δεδομένων. Αν τα σημεία έχουν χρωματική κωδικοποίηση, μπορεί να εμφανιστεί μια επιπλέον μεταβλητή. Τα δεδομένα παρουσιάζονται ως μια συλλογή σημείων, η κάθε μία από τις οποίες έχει την τιμή μίας μεταβλητής που καθορίζει τη θέση στον οριζόντιο άξονα και την τιμή της άλλης μεταβλητής που καθορίζει τη θέση στον κατακόρυφο άξονα (Γιάκας, 2013).

³ Τα επτά εργαλεία είναι:

- a) Έλεγχος φύλλου
- b) Χάρτης ελέγχου
- c) Διαστρωμάτωση(εναλλακτικά, διάγραμμα ροής ή διάγραμμα λειτουργίας)
- d) Διάγραμμα Pareto
- e) Ιστόγραμμα
- f) Διάγραμμα αιτίου-αποτελέσματος (επίσης γνωστό ως διάγραμμα ψαριού ή διάγραμμα Ishikawa)
- g) Διάγραμμα διασποράς

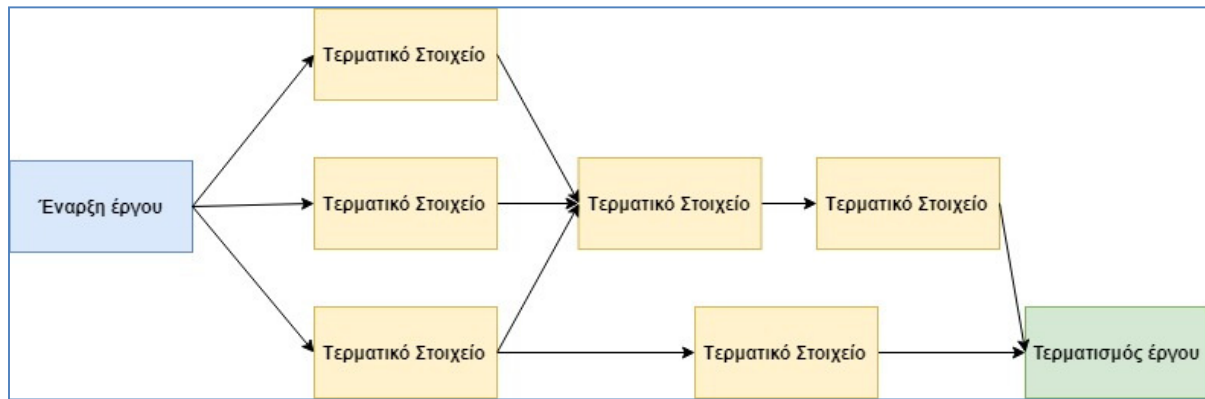


Σχήμα 3.3: Παράδειγμα Διάγραμμα διασποράς.

Μια γραφική παράσταση σκέδασης μπορεί να χρησιμοποιηθεί είτε όταν μία συνεχής μεταβλητή είναι υπό τον έλεγχο του πειραματιστή και η άλλη εξαρτάται από αυτήν είτε όταν και οι δύο συνεχείς μεταβλητές είναι ανεξάρτητες. Εάν υπάρχει μια παράμετρος η οποία αυξάνεται συστηματικά ή / και μειώνεται από την άλλη, ονομάζεται παράμετρος ελέγχου ή ανεξάρτητη μεταβλητή και είναι συνήθως γραφική παράσταση κατά μήκος του οριζόντιου άξονα. Η μετρούμενη ή εξαρτώμενη μεταβλητή συνήθως σχεδιάζεται κατά μήκος του κατακόρυφου άξονα. Εάν δεν υπάρχει εξαρτώμενη μεταβλητή, ο κάθε τύπος μεταβλητής μπορεί να γραφεί σε κάθε άξονα και μια γραφική παράσταση σκέδασης θα απεικονίσει μόνο το βαθμό συσχέτισης (όχι αιτιώδους συνάφειας) μεταξύ δύο μεταβλητών (Χλουβεράκης, 2012).

3.8.4 Διάγραμμα έργου

Ένα Διάγραμμα (Πρόγραμμα) έργου είναι μία γραφική παράσταση (σταθμισμένο κατευθυνόμενο γράφημα) που απεικονίζει την αλληλουχία με την οποία ένα σχέδιο τερματικών στοιχείων πρόκειται να ολοκληρωθούν δείχνοντας τερματικά στοιχεία και τις εξαρτήσεις. Συγκεντρώνεται πάντοτε από αριστερά προς τα δεξιά για να αντικατοπτρίζει τη χρονολογική σειρά του έργου (Χλουβεράκης, 2012).



Σχήμα 3.4: Παράδειγμα Προγράμματος έργου.

Η δομή κατανομής εργασιών ή η δομή κατάρρευσης του προϊόντος παρουσιάζουν τις σχέσεις «εν μέρει». Αντίθετα, το δίκτυο του έργου παρουσιάζει τις σχέσεις «πριν και μετά».

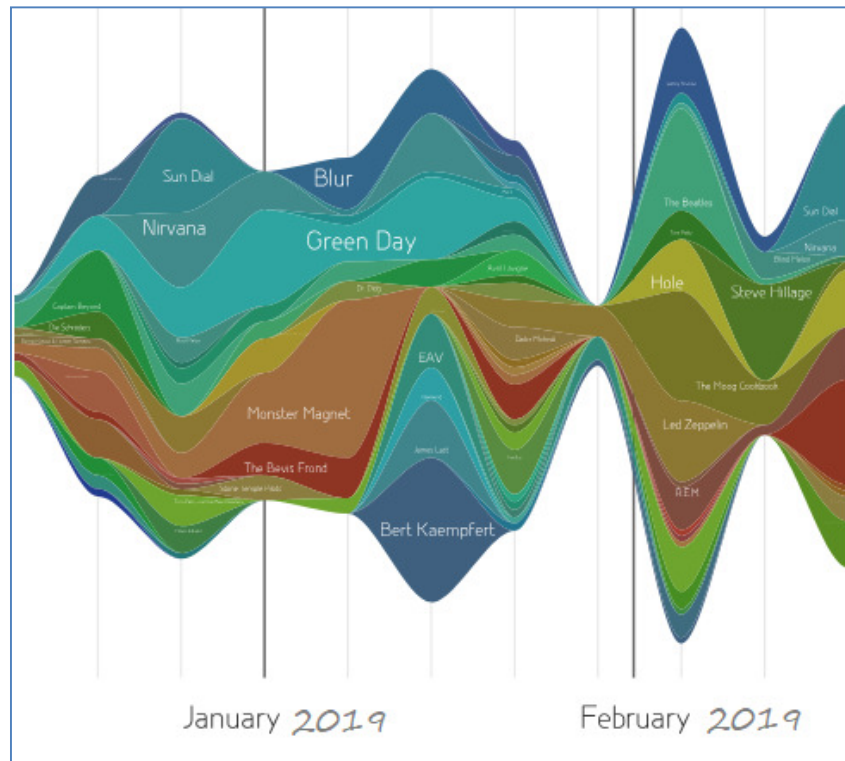
Η προϋπόθεση για ένα έγκυρο δίκτυο έργου είναι ότι δεν περιέχει κυκλικές αναφορές. Οι εξαρτήσεις έργου μπορούν επίσης να απεικονιστούν από έναν πίνακα προκατόχων. Αν και μια τέτοια μορφή είναι πολύ ενοχλητική για την ανθρώπινη ανάλυση, το λογισμικό διαχείρισης έργων συχνά προσφέρει μια τέτοια άποψη για την εισαγωγή δεδομένων.

Ένας εναλλακτικός τρόπος εμφάνισης και ανάλυσης της ακολουθίας των εργασιών του έργου είναι ο πίνακας δομών σχεδιασμού (Χλουβεράκης, 2012).

3.8.5 Streamgraph

Το streamgraph είναι ένας τύπος γραφήματος περιοχής το οποίο μετατοπίζεται γύρω από έναν κεντρικό άξονα, καταλήγοντας σε ένα ρέον, οργανικό σχήμα.

Τα Streamgraphs και η χρήση τους διαδόθηκαν από τον Lee Byron σε ένα άρθρο του περιοδικού New York Times του Φεβρουαρίου 2008 για τα έσοδα του box office (Σιάρδος, 2017).



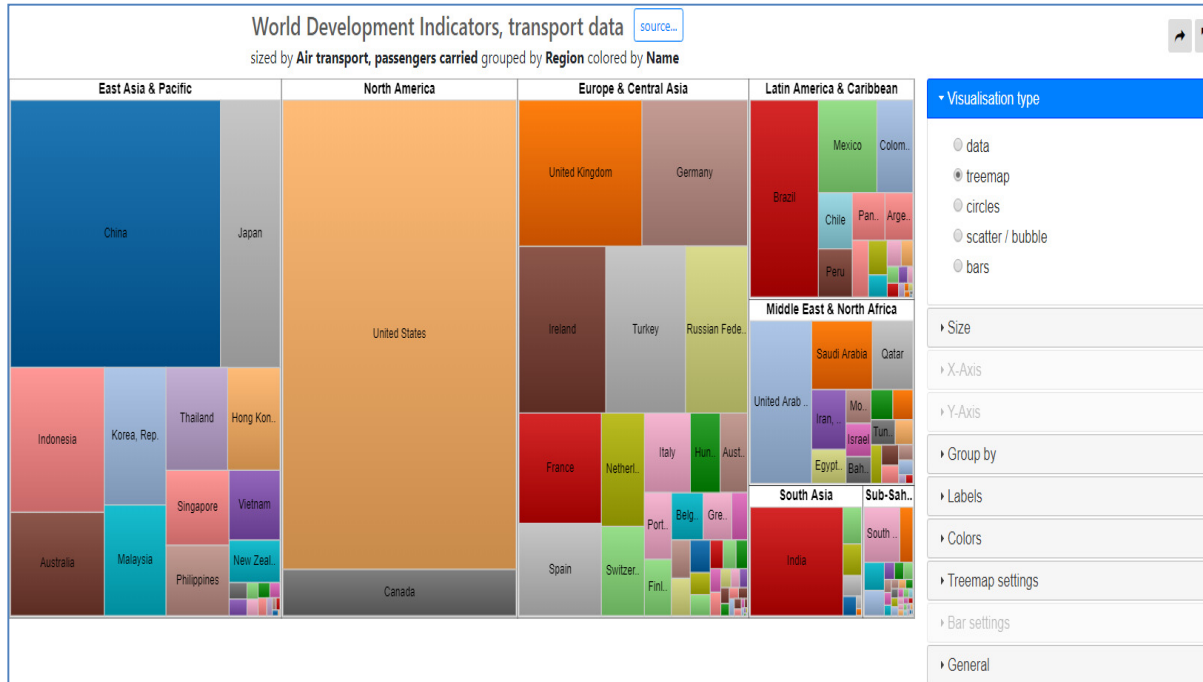
Σχήμα 3.5: Παράδειγμα streamgraph.

3.8.6 Treemapping

Στην απεικόνιση πληροφοριών και στον υπολογισμό, το treemapping είναι μια μέθοδος για την εμφάνιση ιεραρχικών δεδομένων χρησιμοποιώντας ένθετα στοιχεία, συνήθως ορθογώνια (DrasticData - treemap, 2018).

Τα Treemaps εμφανίζουν, ιεραρχικά, δεδομένα ως σύνολο ορθογώνιων κενού. Κάθε κλαδί δέντρου δίνεται σε ένα ορθογώνιο, το οποίο στη συνέχεια είναι πλακιδωτό με μικρότερα ορθογώνια που αντιπροσωπεύουν τα υποκλάδια. Το ορθογώνιο ενός κόμβου φύλλου έχει μια περιοχή ανάλογη προς μια καθορισμένη διάσταση των δεδομένων. Συχνά οι κόμβοι των φύλλων είναι έγχρωμοι για να εμφανίσουν ξεχωριστή διάσταση των δεδομένων (DrasticData - treemap, 2018).

Όταν οι διαστάσεις χρώματος και μεγέθους συσχετίζονται κατά κάποιο τρόπο με τη δομή του δέντρου, συχνά μπορεί κανείς να δει εύκολα σχέδια που θα ήταν δύσκολο να εντοπιστούν με άλλους τρόπους, όπως για παράδειγμα εάν ένα συγκεκριμένο χρώμα είναι ιδιαίτερα σημαντικό. Ένα δεύτερο πλεονέκτημα των treemaps είναι ότι, κατά την κατασκευή, κάνουν αποδοτική χρήση του χώρου. Ως αποτέλεσμα, μπορούν να εμφανίζουν ευανάγνωστα χιλιάδες αντικείμενα στην οθόνη ταυτόχρονα (Σιάρδος, 2017).



Σχήμα 3.6: Παράδειγμα απεικόνισης πληροφοριών treemapping.

(Παγκόσμιος Δείκτης Ανάπτυξης, δεδομένα μεταφορών).

Πηγή: (DrasticData - treemap, 2018).

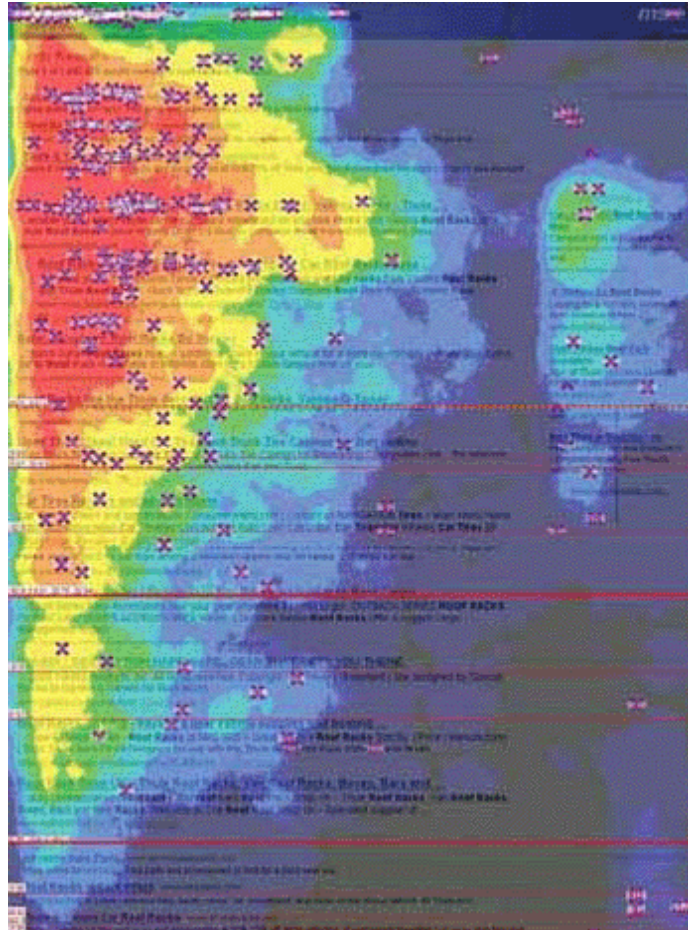
3.8.7 Heatmap

Οι χάρτες θερμικής κατανομής (heatmap) είναι μια γραφική παράσταση των δεδομένων όπου οι μεμονωμένες τιμές που περιέχονται σε μια μήτρα αντιπροσωπεύονται ως χρώματα. Ο «χάρτης θερμικής κατανομής» είναι ένας νεότερος όρος, αλλά οι σκιασμένες μήτρες υπήρχαν εδώ και πάνω από έναν αιώνα (Jaan-Matti, 2013).

Οι χάρτες θερμότητας προέρχονται από 2D απεικονίσεις των τιμών σε ένα πίνακα δεδομένων. Οι μεγαλύτερες τιμές αντιπροσωπεύονται από μικρά σκούρα γκρι ή μαύρα τετράγωνα (εικονοστοιχεία) και οι μικρότερες τιμές από πιο ανοιχτόχρωμα τετράγωνα. Ο Loua (1873) χρησιμοποίησε έναν πίνακα σκίασης για την απεικόνιση των κοινωνικών στατιστικών σε όλες τις περιοχές του Παρισιού. Ο Sneath (1957) εμφάνισε τα αποτελέσματα μιας ανάλυσης συμπλέγματος με την παραλλαγή των σειρών και των στηλών μίας μήτρας για να τοποθετούν παρόμοιες τιμές η μια κοντά στην άλλη σύμφωνα με την ομαδοποίηση. Ο Jacques Bertin χρησιμοποίησε μια παρόμοια αναπαράσταση για την εμφάνιση δεδομένων που αντιστοιχούσαν σε μια κλίμακα Guttman. Η ιδέα να ενταχθούν τα δέντρα συστάδων στις σειρές και τις στήλες του πίνακα δεδομένων προήλθε από τον Robert Ling το 1973. Το Ling χρησιμοποίησε χαρακτηριστικές εκτυπωτών υπερπλήρωσης για να αντιπροσωπεύει διαφορετικές

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

αποχρώσεις του γκρι, ένα πλάτος χαρακτήρων ανά εικονοστοιχείο. Η Leland Wilkinson ανέπτυξε το πρώτο πρόγραμμα υπολογιστή το 1994 (SYSTAT) για την παραγωγή χαρτών θερμικής κατανομής συμπλέγματος με έγχρωμα γραφικά υψηλής ανάλυσης (Jaan-Matti, 2013).



Σχήμα 3.7: Παράδειγμα χάρτη θερμικής κατανομής (heatmap).

Πηγή: (Ampercent - Heatmap, 2018).

3.8.8 Διαγράμματα Spider

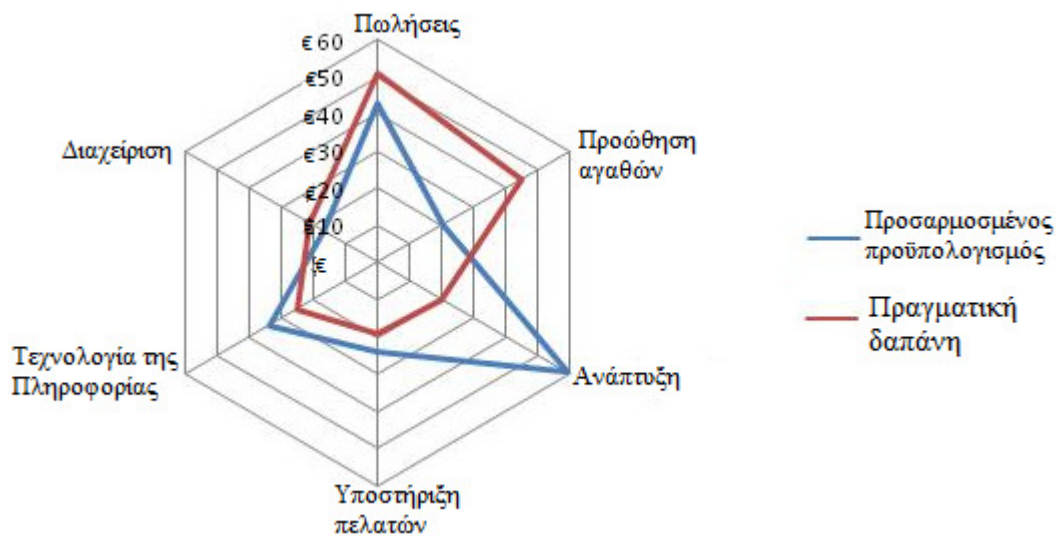
Ένα διάγραμμα Spider είναι μια γραφική μέθοδος εμφάνισης δεδομένων πολλαπλών μεταβλητών με τη μορφή ενός δισδιάστατου χάρτη τριών ή περισσότερων ποσοτικών μεταβλητών που αναπαρίστανται σε άξονες ξεκινώντας από το ίδιο σημείο. Η σχετική θέση και η γωνία των αξόνων είναι τυπικά μη ενημερωτική.

Είναι ισοδύναμο με ένα παράλληλο γράφημα συντεταγμένων, με τους άξονες να είναι διατεταγμένοι ακτινικά. Αποτελείται από μια σειρά από ισόγλωσσες ακτίνες, που ονομάζονται ακτίνες, με κάθε ακτίνα να αντιπροσωπεύει μία από τις μεταβλητές. Το μήκος δεδομένων μιας ακτίνας είναι ανάλογο του μεγέθους της μεταβλητής για το σημείο

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

δεδομένων σε σχέση με το μέγιστο μέγεθος της μεταβλητής σε όλα τα σημεία δεδομένων. Καταρτίζεται μια γραμμή που συνδέει τις τιμές δεδομένων για κάθε ακτίνα. Αυτό δίνει στην επιφάνεια μια αστεροειδή εμφάνιση και την προέλευση ενός από τα δημοφιλή ονόματα για αυτό την επιφάνεια. Η επιφάνεια αστέρι μπορεί να χρησιμοποιηθεί για να απαντήσει στις ακόλουθες ερωτήσεις:

- Ποιες παρατηρήσεις είναι οι περισσότερες παρόμοιες, δηλαδή, υπάρχουν συγκεντρώσεις παρατηρήσεων;
- Υπάρχουν αποθέματα;



Σχήμα 3.8: Αυτό το διάγραμμα Spider αντιπροσωπεύει το διαθέσιμο προϋπολογισμό σε σχέση με τις πραγματικές δαπάνες για έναν συγκεκριμένο οικονομικό οργανισμό.

Τα διαγράμματα Spider είναι ένας χρήσιμος τρόπος για την εμφάνιση πολυμεταβλητών παρατηρήσεων με έναν αυθαίρετο αριθμό μεταβλητών. Κάθε αστέρι αντιπροσωπεύει μία και μόνη παρατήρηση. Συνήθως, οι χάρτες ραντάρ παράγονται σε μορφή πολλαπλών ορίων, με πολλά αστέρια σε κάθε σελίδα και κάθε αστέρι αντιπροσωπεύει μία παρατήρηση.

3.9 Αρχιτεκτονική παρουσίασης δεδομένων

Η αρχιτεκτονική παρουσίασης δεδομένων (Data presentation architecture / DPA) είναι ένα σύνολο δεξιοτήτων που επιδιώκει να εντοπίσει, να χειριστεί, να μορφοποιήσει και να παρουσιάσει δεδομένα με τέτοιο τρόπο ώστε να μεταδώσει με τον καλύτερο δυνατό τρόπο τη σημασία και την ορθή γνώση (Σιάρδος, 2017).

Ιστορικά, ο όρος αρχιτεκτονική παρουσίαση των δεδομένων αποδίδεται από την Kelly Latt (Σιάρδος, 2017): «Η Αρχιτεκτονική Παρουσίαση Δεδομένων (DPA) σπάνια εφαρμόζεται

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

σε ένα σύνολο ικανοτήτων κρίσιμο για την επιτυχία και την αξία της Επιχειρηματικής Ευφυΐας. Η αρχιτεκτονική παρουσίαση των δεδομένων συνδυάζει την επιστήμη των αριθμών, των δεδομένων και τα στατιστικά στοιχεία για την ανακάλυψη πολύτιμων πληροφοριών από τα δεδομένα και την αξιοποίησή τους, σχετικά και εφαρμόσιμα με τις τεχνικές οπτικοποίησης δεδομένων, επικοινωνιών, οργανωτικής ψυχολογίας και διαχείρισης αλλαγών, ώστε να παρέχουν λύσεις επιχειρηματικής ευφυΐας με το εύρος δεδομένων, το χρονοδιάγραμμα παράδοσης, τη μορφή και τις απεικονίσεις που θα υποστηρίζουν και θα οδηγούν πιο αποτελεσματικά την επιχειρησιακή, τακτική και στρατηγική συμπεριφορά προς τους κατανοητούς επιχειρηματικούς (ή οργανωτικούς) στόχους. Η DPA δεν είναι ούτε τεχνολογία πληροφορικής ούτε επιχειρηματική δεξιότητα αλλά υπάρχει ως ξεχωριστό πεδίο εξειδίκευσης. Συχνά συγχέεται με την οπτικοποίηση των δεδομένων, αλλά είναι ένα πολύ ευρύτερο σύνολο δεξιοτήτων που περιλαμβάνει τον προσδιορισμό των δεδομένων σε ποιο χρονοδιάγραμμα και σε ποια ακριβή μορφή πρόκειται να παρουσιαστεί, όχι μόνο ο καλύτερος τρόπος παρουσίασης των δεδομένων που έχουν ήδη επιλεγεί. Οι δεξιότητες οπτικοποίησης δεδομένων είναι ένα στοιχείο της DPA».

Η DPA έχει δύο κύριους στόχους:

- a) Τη χρήση δεδομένων για την παροχή της γνώσης με τον πιο αποτελεσματικό τρόπο (ελαχιστοποιώντας τον θόρυβο, την πολυπλοκότητα και τα άχρηστα δεδομένα ή τις λεπτομέρειες, λαμβάνοντας υπόψη τις ανάγκες και τους ρόλους κάθε ακροατηρίου)
- b) Τη χρήση δεδομένων για την παροχή της γνώσης με τον αποτελεσματικότερο δυνατό τρόπο (παρέχοντας σχετικά, έγκαιρα και πλήρη δεδομένα σε κάθε μέλος του κοινού με σαφή και κατανοητό τρόπο που μεταφέρει σημαντικό νόημα, μπορεί να ασκηθεί προσφυγή και μπορεί να επηρεάσει την κατανόηση, τη συμπεριφορά και τις αποφάσεις).

Με βάση τους παραπάνω στόχους, η πραγματική εργασία της αρχιτεκτονικής παρουσίασης δεδομένων αποτελείται από (Σιάρδος, 2017):

- Δημιουργία αποτελεσματικών μηχανισμών προβολής για κάθε μέλος του ακροατηρίου, ανάλογα με το ρόλο, τα καθήκοντά τους, τις τοποθεσίες τους και την πρόσβαση στην τεχνολογία
- Ορίζοντα με σημαντικό νόημα (σχετική γνώση) που απαιτείται από κάθε μέλος του ακροατηρίου σε κάθε πλαίσιο
- Προσδιορισμό της απαιτούμενης περιοδικότητας των ενημερώσεων δεδομένων (το νόμισμα των δεδομένων)

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

- Προσδιορισμό του σωστού χρονισμού για την παρουσίαση δεδομένων (πότε και πόσος χρόνος χρειάζεται ο χρήστης για να δει τα δεδομένα)
- Εύρεση των σωστών δεδομένων (περιοχή θέματος, ιστορική προσέγγιση, εύρος, επίπεδο λεπτομέρειας κ.λπ.)
- Χρήση κατάλληλης ανάλυσης, ομαδοποίησης, οπτικοποίησης και άλλων μορφών παρουσίασης

Το έργο της DPA έχει κοινά σημεία με διάφορους άλλους τομείς, όπως (Σιάρδος, 2017):

- Επιχειρηματική ανάλυση για τον καθορισμό των επιχειρηματικών στόχων, των απαιτήσεων συλλογής, των διαδικασιών χαρτογράφησης.
- Βελτίωση επιχειρηματικών διαδικασιών στο βαθμό που ο στόχος της είναι η βελτίωση και ο εξορθολογισμός των ενεργειών και των αποφάσεων για την προώθηση των επιχειρηματικών στόχων
- Οπτικοποίηση δεδομένων, δεδομένου ότι χρησιμοποιεί καθιερωμένες θεωρίες οπτικοποίησης για να προσθέσει ή να τονίσει τη σημασία στην παρουσίαση δεδομένων
- Αρχιτεκτονική των πληροφοριών, αλλά η αρχιτεκτονική της πληροφορίας επικεντρώνεται στα μη δομημένα δεδομένα και συνεπώς αποκλείει τόσο την ανάλυση (από στατιστική άποψη) όσο και την άμεση μετατροπή του πραγματικού περιεχομένου (δεδομένα για την DPA) σε νέες οντότητες και συνδυασμούς.
- HCI και σχεδιασμός αλληλεπίδρασης, καθώς πολλές από τις αρχές για τον τρόπο σχεδιασμού της απεικόνισης διαδραστικών δεδομένων έχουν αναπτυχθεί διεπιστημονικά με το HCI.
- Οπτική δημοσιογραφία που βασίζεται στα δεδομένα της δημοσιογραφίας. Η Visual δημοσιογραφία ασχολείται με όλα τα είδη των γραφικών διευκόλυνσης της αφήγησης των ειδήσεων. Παρόλα αυτά, ο τομέας της δημοσιογραφίας βρίσκεται στην πρώτη γραμμή στην ανάπτυξη νέων οπτικοποιήσεων δεδομένων για την επικοινωνία των δεδομένων.
- Γραφιστική σχεδίαση, μεταφορά πληροφοριών μέσω τυπογραφίας, θέσης και άλλων αισθητικών ανησυχιών.

4 Κεφάλαιο: «Επιχειρηματική ευφυΐα»

4.1 Λειτουργίες και τεχνολογίες επιχειρηματικής ευφυΐας

Η Επιχειρηματική ευφυΐα (Business Intelligence / BI) περιλαμβάνει τις στρατηγικές και τις τεχνολογίες που χρησιμοποιούν οι επιχειρήσεις για την ανάλυση δεδομένων των επιχειρηματικών πληροφοριών (είναι ένα από τα τρία βασικά τμήματα της βιομηχανίας πληροφοριών). Οι τεχνολογίες Επιχειρηματικής ευφυΐας προσφέρουν ιστορικές, τρέχουσες και προγνωστικές απόψεις των επιχειρήσεων. Οι κοινές λειτουργίες των τεχνολογιών επιχειρηματικής ευφυΐας είναι οι εξής (Κύρκος, 2015):

- **Επιχειρηματικές αναφορές**, αφορούν τόσο τη «δημόσια αναφορά λειτουργικών και χρηματοοικονομικών δεδομένων από μια επιχείρηση» όσο και την «τακτική παροχή πληροφοριών στους υπεύθυνους για τη λήψη αποφάσεων εντός ενός οργανισμού για την υποστήριξη τους στην εργασία τους». Αποτελεί θεμελιώδες μέρος της μεγαλύτερης κίνησης προς τη βελτίωση της επιχειρησιακής νοημοσύνης και της διαχείρισης της γνώσης. Η εφαρμογή συχνά περιλαμβάνει διαδικασίες εξαγωγής, μετασχηματισμού και φόρτωσης σε συντονισμό με μια αποθήκη δεδομένων και στη συνέχεια χρησιμοποιώντας ένα ή περισσότερα εργαλεία αναφοράς. Οι αναφορές μπορούν να διανεμηθούν σε έντυπη μορφή, μέσω ηλεκτρονικού ταχυδρομείου ή μέσω εταιρικού ενδοδικτύου.
- **Ηλεκτρονική αναλυτική επεξεργασία**, (Online analytical processing / OLAP), είναι μια προσέγγιση για την ταχεία απάντηση σε ερωτήματα πολυδιάστατης ανάλυσης (multidimensional analysis / MDA) στον υπολογισμό. Το OLAP αποτελεί μέρος της ευρύτερης κατηγορίας επιχειρηματικών πληροφοριών, η οποία περιλαμβάνει επίσης σχεσιακές βάσεις δεδομένων, γραπτή αναφορά και εξόρυξη δεδομένων. Οι τυπικές εφαρμογές του OLAP περιλαμβάνουν επιχειρηματικές αναφορές για πωλήσεις, μάρκετινγκ, αναφορά διαχείρισης, διαχείριση επιχειρησιακών διαδικασιών (Business process management / BPM), προϋπολογισμό και πρόβλεψη, χρηματοοικονομικές εκθέσεις και παρόμοιες περιοχές, με νέες εφαρμογές.
- **Analytics**, είναι η ανακάλυψη, η ερμηνεία και η επικοινωνία σημαντικών προτύπων στα δεδομένα και τη διαδικασία εφαρμογής αυτών των προτύπων για την αποτελεσματική λήψη αποφάσεων. Με άλλα λόγια, τα αναλυτικά στοιχεία μπορούν να κατανοηθούν ως ο συνδετικός ιστός μεταξύ δεδομένων και αποτελεσματικής λήψης αποφάσεων, μέσα σε έναν οργανισμό.

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

- **Εξόρυξη δεδομένων**, είναι η διαδικασία της ανεύρεσης σχεδίων σε μεγάλα σύνολα δεδομένων που περιλαμβάνουν μεθόδους στη διασταύρωση της μηχανικής μάθησης, των στατιστικών στοιχείων και των συστημάτων βάσης δεδομένων.
- **Διαδικασία Εξόρυξης**, είναι μια οικογένεια τεχνικών στον τομέα της διαχείρισης της διαδικασίας που υποστηρίζουν την ανάλυση των επιχειρηματικών διαδικασιών με βάση τα αρχεία καταγραφής συμβάντων. Κατά την εξόρυξη διεργασιών, εξειδικευμένοι αλγόριθμοι εξόρυξης δεδομένων εφαρμόζονται στα δεδομένα του ημερολογίου συμβάντων προκειμένου να εντοπιστούν οι τάσεις, τα πρότυπα και οι λεπτομέρειες που περιέχονται στα αρχεία καταγραφής συμβάντων που καταγράφονται από ένα σύστημα πληροφοριών.
- **Σύνθετη επεξεργασία συμβάντων**, είναι μια μέθοδος ανίχνευσης και ανάλυσης (επεξεργασίας) ροών πληροφοριών (δεδομένων) σχετικά με τα πράγματα που συμβαίνουν (γεγονότα) και την εξαγωγή συμπερασμάτων από αυτά. Η σύνθετη επεξεργασία συμβάντων (Complex event processing / CEP) είναι η επεξεργασία συμβάντων που συνδυάζει δεδομένα από πολλαπλές πηγές για να συναγάγει γεγονότα ή μοτίβα που υποδεικνύουν πιο περίπλοκες περιστάσεις.
- **Διαχείριση των επιδόσεων των επιχειρήσεων**, είναι ένα σύνολο διαδικασιών διαχείρισης απόδοσης και αναλύσεων που επιτρέπει τη διαχείριση της απόδοσης ενός οργανισμού για την επίτευξη ενός ή περισσότερων προεπιλεγμένων στόχων.
- **Συγκριτική αξιολόγηση**, είναι η πρακτική της σύγκρισης των επιχειρηματικών διαδικασιών και των μετρήσεων απόδοσης με τις καλύτερες πρακτικές και τις βέλτιστες πρακτικές άλλων εταιρειών. Οι διαστάσεις που συνήθως μετριοούνται είναι η ποιότητα, ο χρόνος και το κόστος.
- **Εξόρυξη κειμένου**, είναι η διαδικασία απόκτησης υψηλής ποιότητας πληροφοριών από το κείμενο. Οι υψηλής ποιότητας πληροφορίες προέρχονται συνήθως από την εκπόνηση προτύπων και τάσεων μέσω μέσων όπως η στατιστική μάθηση των προτύπων. Η εξόρυξη κειμένου συνήθως περιλαμβάνει τη διαδικασία διαμόρφωσης του κειμένου εισαγωγής, την παραγωγή μοτίβων μέσα στα δομημένα δεδομένα και τέλος την αξιολόγηση και ερμηνεία της παραγωγής. Η «υψηλή ποιότητα» στην εξόρυξη κειμένου συνήθως αναφέρεται σε κάποιο συνδυασμό συνάφειας, καινοτομίας και ενδιαφέροντος. Οι τυπικές εργασίες εξόρυξης κειμένου περιλαμβάνουν την κατηγοριοποίηση κειμένου, την ομαδοποίηση κειμένου, την εξαγωγή ιδεών / οντοτήτων,

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

την ανάλυση συναισθημάτων, τη συνοπτική παρουσίαση εγγράφων και τη μοντελοποίηση σχέσεων οντοτήτων.

- **Προγνωστικές αναλύσεις**, περιλαμβάνει μια ποικιλία στατιστικών τεχνικών από την εξόρυξη δεδομένων, τη μέθοδο πρόβλεψης και τη μηχανική μάθηση, που αναλύει τα τρέχοντα και τα ιστορικά δεδομένα για να κάνουν προβλέψεις για μελλοντικά ή άγνωστα συμβάντα.
- **Συντακτικές αναλύσεις**, είναι η τρίτη και τελική φάση της επιχειρηματικής ανάλυσης η οποία περιλαμβάνει επίσης περιγραφικά και προγνωστικά στοιχεία ανάλυσης.

Οι τεχνολογίες BI μπορούν να χειριστούν μεγάλα ποσά δομημένων και μερικές φορές αδόμητων δεδομένων για να βοηθήσουν στον εντοπισμό, την ανάπτυξη και με άλλο τρόπο τη δημιουργία νέων στρατηγικών επιχειρηματικών ευκαιριών. Σκοπός τους είναι να επιτρέπουν την εύκολη ερμηνεία αυτών των μεγάλων δεδομένων. Ο προσδιορισμός νέων ευκαιριών και η εφαρμογή μιας αποτελεσματικής στρατηγικής βασισμένης σε πληροφορίες μπορεί να προσφέρει στις επιχειρήσεις ένα ανταγωνιστικό πλεονέκτημα στην αγορά και μακροπρόθεσμη σταθερότητα (Σταλίδης & Καρδαράς, 2015).

Οι επιχειρησιακές μυστικές υπηρεσίες μπορούν να χρησιμοποιηθούν από τις επιχειρήσεις για να υποστηρίξουν ένα ευρύ φάσμα επιχειρηματικών αποφάσεων που κυμαίνονται από λειτουργικές σε στρατηγικές. Οι βασικές λειτουργικές αποφάσεις περιλαμβάνουν την τοποθέτηση ή την τιμολόγηση του προϊόντος. Οι στρατηγικές επιχειρηματικές αποφάσεις περιλαμβάνουν προτεραιότητες, στόχους και κατευθύνσεις στο ευρύτερο επίπεδο. Σε όλες τις περιπτώσεις, η BI είναι πιο αποτελεσματική όταν συνδυάζει δεδομένα που προέρχονται από την αγορά στην οποία λειτουργεί μια εταιρεία (εξωτερικά δεδομένα) με δεδομένα από εσωτερικές πηγές της επιχείρησης όπως τα οικονομικά και τα λειτουργικά δεδομένα (εσωτερικά δεδομένα). Όταν συνδυάζονται, τα εξωτερικά και τα εσωτερικά δεδομένα μπορούν να παράσχουν μια πλήρη εικόνα η οποία, στην πραγματικότητα, δημιουργεί μια «νοημοσύνη» που δεν μπορεί να προέλθει από κάθε μοναδικό σύνολο δεδομένων (Σταλίδης & Καρδαράς, 2015).

Μεταξύ των μυριάδων χρήσεων, τα εργαλεία επιχειρησιακής νοημοσύνης εξουσιοδοτούν τους οργανισμούς να αποκτήσουν εικόνα για νέες αγορές, να αξιολογήσουν τη ζήτηση και την καταλληλότητα προϊόντων και υπηρεσιών για διάφορα τμήματα της αγοράς και να μετρήσουν τον αντίκτυπο των προσπαθειών μάρκετινγκ.

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

Συχνά οι εφαρμογές BI χρησιμοποιούν δεδομένα που συλλέγονται από μια αποθήκη δεδομένων (data warehouse / DW). Μια αποθήκη δεδομένων περιέχει ένα αντίγραφο αναλυτικών δεδομένων που διευκολύνουν την υποστήριξη λήψης αποφάσεων.

4.2 Χαρακτηριστικά Επιχειρηματικής ευφυΐας

Σύμφωνα με τον George Colony (2016), Διευθύνοντα Σύμβουλο της Forrester Research (αμερικανική εταιρεία έρευνας αγοράς νέων τεχνολογιών), η επιχειρησιακή ευφυΐα είναι ένα σύνολο μεθοδολογιών, διαδικασιών, αρχιτεκτονικών και τεχνολογιών που μετατρέπουν τα ανεπεξέργαστα δεδομένα σε χρήσιμες πληροφορίες που χρησιμοποιούνται για την αποτελεσματικότερη στρατηγική, τακτική και επιχειρησιακή κατανόηση και λήψη αποφάσεων. Σύμφωνα με τον ορισμό αυτό, η επιχειρηματική ευφυΐα περιλαμβάνει τη διαχείριση των πληροφοριών (ενοποίηση των δεδομένων, την ποιότητα των δεδομένων, αποθήκευση δεδομένων, διαχείριση master-δεδομένων, κειμένου και περιεχομένου Analytics). Ως εκ τούτου, ο Colony αναφέρεται στην προετοιμασία δεδομένων και τη χρήση δεδομένων ως δύο χωριστά αλλά στενά συνδεδεμένα τμήματα της αρχιτεκτονικής στοίβας επιχειρηματικών πληροφοριών (Liang & Liu, 2018).

Μερικά στοιχεία επιχειρηματικής ευφυΐας είναι (Liang & Liu, 2018):

- Πολυδιάστατη συσσώρευση και κατανομή
- Αποσυναρμολόγηση, ετικέτες και τυποποίηση
- Αναφορές σε πραγματικό χρόνο με αναλυτική ειδοποίηση
- Μια μέθοδος διασύνδεσης με μη δομημένες πηγές δεδομένων
- Ομαδοποίηση, κατάρτιση προϋπολογισμού και κυλιόμενες προβλέψεις
- Στατιστική εξαγωγή και πιθανή προσομοίωση
- Βελτιστοποίηση βασικών δεικτών απόδοσης
- Έλεγχος έκδοσης και διαχείριση διαδικασίας
- Άνοιγμα διαχείρισης στοιχείων

Ο George Colony διακρίνει αυτά από την αγορά επιχειρηματικών πληροφοριών, η οποία είναι «μόνο τα κορυφαία στρώματα της αρχιτεκτονικής στοίβας BI, όπως η αναφορά, τα αναλυτικά στοιχεία και τα dashboards (πίνακες ελέγχου συχνά παρέχουν γρήγορες απόψεις βασικών δεικτών απόδοσης (key performance indicators / KPIs) που σχετίζονται με συγκεκριμένο στόχο ή επιχειρηματική διαδικασία».

4.2.1 Σύγκριση με την ανταγωνιστική νοημοσύνη

Αν και ο όρος επιχειρησιακή ευφυΐα είναι μερικές φορές συνώνυμο της ανταγωνιστικής νοημοσύνης / Competitive intelligence (είναι η δράση καθορισμού, συλλογής, ανάλυσης και διανομής πληροφοριών σχετικά με προϊόντα, πελάτες, ανταγωνιστές και οποιαδήποτε πτυχή του περιβάλλοντος που απαιτείται για την υποστήριξη στελεχών και διαχειριστών στη λήψη στρατηγικών αποφάσεων για έναν οργανισμό), επειδή και οι δύο υποστηρίζουν τη λήψη αποφάσεων, ωστόσο η ΒΙ χρησιμοποιεί τεχνολογίες, διαδικασίες και εφαρμογές για την ανάλυση κυρίως εσωτερικών, δομημένων δεδομένων και επιχειρηματικών διαδικασιών, ενώ η ανταγωνιστική νοημοσύνη συγκεντρώνει, αναλύει και διαδίδει πληροφορίες, επικεντρώνεται σε ανταγωνιστές της εταιρείας. Εάν γίνει ευρεία κατανόηση, η επιχειρηματική ευφυΐα μπορεί να περιλαμβάνει το υποσύνολο της ανταγωνιστικής νοημοσύνης (Σταλίδης & Καρδαράς, 2015).

4.2.2 Σύγκριση με τις επιχειρηματικές αναλύσεις

Οι επιχειρηματικές πληροφορίες και οι επιχειρηματικές αναλύσεις (αναφέρονται στις δεξιότητες, τις τεχνολογίες, τις πρακτικές για συνεχή επαναληπτική διερεύνηση και διερεύνηση των επιχειρηματικών επιδόσεων του παρελθόντος για να αποκτήσουν γνώση και να οδηγήσουν τον επιχειρηματικό προγραμματισμό) μερικές φορές χρησιμοποιούνται εναλλακτικά, αλλά υπάρχουν εναλλακτικοί ορισμοί. Ο Thomas Davenport, καθηγητής τεχνολογίας πληροφορικής και διαχείρισης στο Babson College υποστηρίζει ότι η επιχειρησιακή ευφυΐα πρέπει να χωριστεί σε ερωτήματα, αναφορές, σε απευθείας σύνδεση αναλυτική επεξεργασία (OLAP), ένα εργαλείο «ειδοποιήσεων» και επιχειρηματικές αναλύσεις. Στον ορισμό αυτό, οι επιχειρηματικές αναλύσεις είναι το υποσύνολο του ΒΙ που επικεντρώνεται στις στατιστικές, την πρόβλεψη και τη βελτιστοποίηση, αντί της λειτουργικότητας αναφοράς (Σταλίδης & Καρδαράς, 2015).

4.3 Στοιχεία της Επιχειρηματικής ευφυΐας

Τα βασικά στοιχεία της Επιχειρηματικής ευφυΐας αντικατοπτρίζουν ένα περίπλοκο σύστημα το οποίο περνάει τα δεδομένα για να μετατραπούν σε πληροφορίες. Ένα από τα πρώτα βήματα για την εκκίνηση ενός προγράμματος ΒΙ, είναι η κατανόηση όλων των συστατικών αυτού του τοπίου. Οι ιδιαιτερότητες αυτού του συστήματος τείνουν να διαφέρουν ανάλογα με τη βιομηχανία και την οργάνωση, αλλά σε μακροοικονομικό επίπεδο,

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

όλα τα τοπία BI έχουν την ίδια μορφή. Συνήθως αποτελείται από πέντε πυλώνες και πέντε μπλοκ θεμελίωσης (Σταλιδής & Καρδαράς, 2015):

Οι πέντε πυλώνες:

1. Πηγές δεδομένων
2. Ενσωμάτωση δεδομένων
3. Διαχείριση δεδομένων
4. Αναφορές
5. Διάδοση πληροφοριών

Τα πέντε σημεία θεμελίωσης:

1. Ασφάλεια πληροφοριών
2. Ποιότητα δεδομένων
3. Διαχείριση μεταδεδομένων
4. Διακυβέρνηση δεδομένων
5. Άνθρωποι και πολιτισμός

4.4 Δεδομένα επιχειρηματικών λειτουργιών

Οι επιχειρηματικές λειτουργίες μπορούν να παράγουν ένα πολύ μεγάλο αριθμό πληροφοριών, όπως μηνύματα ηλεκτρονικού ταχυδρομείου, σημειώματα, σημειώσεις από κέντρα κλήσεων, ειδήσεις, ομάδες χρηστών, συζητήσεις, αναφορές, ιστοσελίδες, παρουσιάσεις, αρχεία εικόνας, αρχεία βίντεο και υλικό μάρκετινγκ. Σύμφωνα με τον Daniel Sewell (2018), πάνω από το 85% όλων των επιχειρηματικών πληροφοριών υπάρχει σε αυτές τις μορφές. Μια εταιρεία μπορεί να χρησιμοποιήσει ένα τέτοιο έγγραφο μόνο μία φορά. Λόγω του τρόπου με τον οποίο παράγεται και αποθηκεύεται, οι πληροφορίες αυτές είναι είτε μη δομημένες (δηλαδή πληροφορίες που είτε δεν έχουν προκαθορισμένο μοντέλο δεδομένων είτε δεν είναι οργανωμένες με προκαθορισμένο τρόπο) είτε ημιδομημένες (δηλαδή μια μορφή δομημένων δεδομένων που δεν συμμορφώνονται με την τυπική δομή των μοντέλων δεδομένων που σχετίζονται με σχεσιακές βάσεις δεδομένων ή άλλες μορφές πίνακα δεδομένων, αλλά περιέχουν ετικέτες ή άλλους δείκτες για να διαχωρίσουν σημασιολογικά στοιχεία και να επιβάλουν ιεραρχίες αρχεία και πεδία εντός των δεδομένων) (Sewell, 2018).

Η διαχείριση ημιδομημένων δεδομένων είναι ένα άλυτο πρόβλημα στον κλάδο της τεχνολογίας των πληροφοριών. Σύμφωνα με τις προβλέψεις της Gartner (2003), οι εργαζόμενοι περνούν το 30-40% του χρόνου τους αναζητώντας, βρίσκοντας και

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

αξιολογώντας μη δομημένα δεδομένα. Το BI χρησιμοποιεί τόσο δομημένα όσο και μη δομημένα δεδομένα. Το πρώτο είναι εύκολο να αναζητηθεί και το τελευταίο περιέχει μια μεγάλη ποσότητα των πληροφοριών που απαιτούνται για την ανάλυση και τη λήψη αποφάσεων. Λόγω της δυσκολίας της σωστής αναζήτησης, εύρεσης και αξιολόγησης μη δομημένων ή ημιδομημένων δεδομένων, οι οργανώσεις δεν μπορούν να αντλήσουν από αυτές τις τεράστιες δεξαμενές πληροφοριών, οι οποίες θα μπορούσαν να επηρεάσουν μια συγκεκριμένη απόφαση ή έργο. Αυτό μπορεί τελικά να οδηγήσει σε άσχημη λήψη αποφάσεων (Sewell, 2018).

Επομένως, κατά το σχεδιασμό μιας επιχειρησιακής νοημοσύνης / DW-λύσης, πρέπει να αντιμετωπιστούν τα συγκεκριμένα προβλήματα που σχετίζονται με ημιδομημένα και μη δομημένα δεδομένα, καθώς και τα ειδικά για τα δομημένα δεδομένα.

4.4.1 Μη δομημένα δεδομένα έναντι ημιδομημένων δεδομένων

Τα μη δομημένα και ημιδομημένα δεδομένα έχουν διαφορετικές σημασίες ανάλογα με το περιβάλλον τους. Στο πλαίσιο συστημάτων σχεσιακών βάσεων δεδομένων, τα μη δομημένα δεδομένα δεν μπορούν να αποθηκευτούν σε προβλέψιμες διατεταγμένες στήλες και σειρές. Ένας τύπος μη δομημένων δεδομένων τυπικά αποθηκεύεται σε ένα BLOB (δυαδικό μεγάλο αντικείμενο / Binary Large Object), ένας τύπος δεδομένων τύπου catch-all που είναι διαθέσιμος στα περισσότερα συστήματα διαχείρισης σχεσιακών βάσεων δεδομένων (είναι μια ψηφιακή βάση δεδομένων βασισμένη στο σχεσιακό μοντέλο δεδομένων). Τα μη δομημένα δεδομένα μπορούν επίσης να αναφέρονται σε ακανόνιστα ή τυχαία επαναλαμβανόμενα μοτίβα στηλών που διαφέρουν από σειρά σε σειρά ή αρχεία φυσικής γλώσσας που δεν έχουν λεπτομερή μεταδεδομένα.

Ωστόσο, πολλοί από αυτούς τους τύπους δεδομένων, όπως τα ηλεκτρονικά μηνύματα, τα αρχεία κειμένου επεξεργασίας κειμένου, τα αρχεία PPT, τα αρχεία εικόνας και τα αρχεία βίντεο, συμμορφώνονται με ένα πρότυπο που προσφέρει τη δυνατότητα μεταδεδομένων. Τα μεταδεδομένα μπορούν να περιλαμβάνουν πληροφορίες όπως ο δημιουργός και ο χρόνος δημιουργίας και αυτό μπορεί να αποθηκευτεί σε σχεσιακή βάση δεδομένων. Επομένως, μπορεί να είναι πιο ακριβές να αναφέρονται αυτά ως ημιδομημένα έγγραφα ή στοιχεία, αλλά δεν φαίνεται να έχει επιτευχθεί συγκεκριμένη συναίνεση.

Τα μη δομημένα δεδομένα μπορούν επίσης να είναι απλά η γνώση που έχουν οι επιχειρηματικοί χρήστες σχετικά με τις μελλοντικές επιχειρηματικές τάσεις. Οι επιχειρηματικές προβλέψεις ευθυγραμμίζονται φυσικά με το σύστημα BI, επειδή οι

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας επιχειρησιακοί χρήστες σκέφτονται την επιχείρησή τους σε συγκεντρωτικούς όρους. Η καταγραφή των επιχειρηματικών γνώσεων που μπορεί να υπάρχουν μόνο στο μυαλό των επιχειρηματικών χρηστών παρέχει μερικά από τα πιο σημαντικά σημεία δεδομένων για μια ολοκληρωμένη λύση BI (Sewell, 2018).

4.4.2 Περιορισμοί ημιδομημένων και αδόμητων δεδομένων

Υπάρχουν διάφορες προκλήσεις για την ανάπτυξη BI με ημι-δομημένα δεδομένα. Σύμφωνα με τον Inmon & Nesavich, μερικά από αυτά είναι (Sewell, 2018):

- Πρόσβαση σε μη δομημένα κείμενα δεδομένων - τα μη δομημένα δεδομένα αποθηκεύονται σε μια τεράστια ποικιλία μορφών.
- Ορολογία - Μεταξύ των ερευνητών και των αναλυτών, υπάρχει ανάγκη να αναπτυχθεί τυποποιημένη ορολογία.
- Όγκος δεδομένων - Όπως αναφέρθηκε προηγουμένως, μέχρι και το 85% όλων των δεδομένων υπάρχουν ως ημιδομημένα δεδομένα.
- Δυνατότητα αναζήτησης μη δομημένων κειμένων δεδομένων - Μια απλή αναζήτηση σε ορισμένα δεδομένα, για παράδειγμα «αυτοκίνητο», οδηγεί σε συνδέσμους όπου υπάρχει αναφορά σε αυτόν τον ακριβή όρο αναζήτησης.

4.5 Μεταδεδομένα

Για την επίλυση προβλημάτων σχετικά με την αναζήτηση και την αξιολόγηση των δεδομένων, είναι απαραίτητο να γνωρίζουμε κάτι σχετικά με το περιεχόμενο. Αυτό μπορεί να γίνει με την προσθήκη πλαισίου μέσω της χρήσης μεταδεδομένων.

Τα μεταδεδομένα είναι δεδομένα (πληροφορίες) που παρέχουν πληροφορίες για άλλα δεδομένα. Υπάρχουν πολλά διαφορετικά είδη μεταδεδομένων, μεταξύ αυτών των περιγραφικών μεταδεδομένων, των διαρθρωτικών μεταδεδομένων, των διοικητικών μεταδεδομένων, των μεταδεδομένων αναφοράς και των στατιστικών μεταδεδομένων. Αναλυτικότερα (Sewell, 2018):

- Τα **περιγραφικά μεταδεδομένα**, περιγράφουν έναν πόρο για σκοπούς όπως η ανακάλυψη και η αναγνώριση. Μπορεί να περιλαμβάνει στοιχεία όπως τίτλο, περίληψη, συγγραφέα και λέξεις-κλειδιά.
- Τα **διαρθρωτικά μεταδεδομένα**, είναι μεταδεδομένα σχετικά με τα δοχεία δεδομένων και υποδεικνύουν τον τρόπο σύνθεσης σύνθετων αντικειμένων, για παράδειγμα, τον

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

τρόπο με τον οποίο οι σελίδες παραγγέλνονται για να σχηματίσουν κεφάλαια. Περιγράφει τους τύπους, τις εκδόσεις, τις σχέσεις και άλλα χαρακτηριστικά των ψηφιακών υλικών.

- Τα **μεταδεδομένα διαχείρισης**, παρέχουν πληροφορίες για τη διαχείριση ενός πόρου, όπως πότε και πώς δημιουργήθηκε, τύπος αρχείου και άλλες τεχνικές πληροφορίες και ποιος μπορεί να έχει πρόσβαση σε αυτόν.
- Τα **μεταδεδομένα αναφοράς**, περιγράφουν το περιεχόμενο και την ποιότητα των στατιστικών δεδομένων.
- Τα **στατιστικά μεταδεδομένα**, μπορούν επίσης να περιγράφουν διαδικασίες που συλλέγουν, επεξεργάζονται ή παράγουν στατιστικά δεδομένα. Αυτά τα μεταδεδομένα καλούνται επίσης δεδομένα διεργασίας.

Πολλά συστήματα ήδη καταγράφουν μερικά μεταδεδομένα (π.χ. όνομα αρχείου, συγγραφέα, μέγεθος κ.λπ.), αλλά πιο χρήσιμα θα ήταν τα μεταδεδομένα για το πραγματικό περιεχόμενο - π.χ. περιλήψεις, θέματα, άτομα ή εταιρείες που αναφέρθηκαν. Δύο τεχνολογίες που έχουν σχεδιαστεί για τη δημιουργία μεταδεδομένων σχετικά με το περιεχόμενο είναι αυτόματη κατηγοριοποίηση και εξαγωγή πληροφοριών.

4.6 Εφαρμογές Επιχειρηματικής ευφυΐας

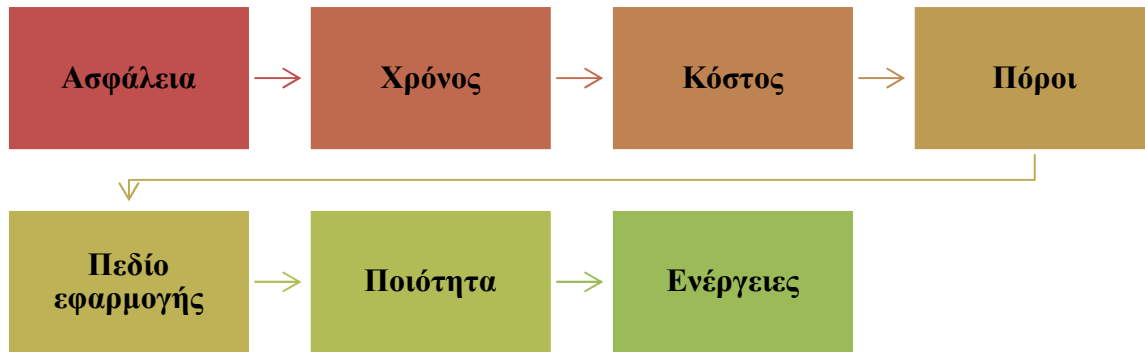
Οι επιχειρηματικές πληροφορίες μπορούν να χρησιμοποιηθούν για τους ακόλουθους επιχειρηματικούς σκοπούς (Κύρκος, 2015):

4.6.1 Μετρήσεις απόδοσης

Μια μέτρηση απόδοσης μετρά τη συμπεριφορά, τις δραστηριότητες και την απόδοση ενός οργανισμού. Αξιολογεί πόσο καλά εργάζονται οι εργαζόμενοι και πώς οι επιχειρήσεις επιτυγχάνουν τους στόχους τους. Παρέχει σκληρά δεδομένα και εκπέμπει αποτελέσματα που εκτιμούν σαφώς προσδιορισμένες ποσότητες σε ένα εύρος που διευκολύνει τη βελτίωση και την αναβάθμιση. Στην ιδανική περίπτωση, οι καλές μετρήσεις απόδοσης αποτελούν τη βάση για την καλύτερη επίτευξη των συνολικών στόχων μιας μικρής επιχείρησης.

Πρέπει να υποστηρίζει μια σειρά από ανάγκες των ενδιαφερομένων από τους πελάτες, τους μετόχους και τους εργαζομένους. Παρόλο που παραδοσιακά οι μετρήσεις βασίζονται στη χρηματοδότηση και εστιάζουν εσωτερικά στην απόδοση του οργανισμού, οι μετρήσεις μπορούν επίσης να επικεντρωθούν στην απόδοση σε σχέση με τις απαιτήσεις και την αξία των πελατών.

Στη διαχείριση έργων, οι μετρήσεις απόδοσης χρησιμοποιούνται για την αξιολόγηση της συνολικής ευημερίας ενός έργου και επικεντρώνονται στην εκτίμησή του βάσει επτά κριτηρίων, σύμφωνα με το παρακάτω Σχήμα 3.1:



Σχήμα 4.1: Κριτήρια εκτίμησης Μέτρησης απόδοσης στην διαχείριση έργων.

Πηγή: (Κύρκος, 2015).

Στα κέντρα κλήσεων, οι μετρήσεις απόδοσης συμβάλλουν στην καταγραφή της εσωτερικής απόδοσης και μπορούν να περιλαμβάνουν μετρήσεις παραγωγικότητας και την ποιότητα των υπηρεσιών που παρέχει ο σύμβουλος εξυπηρέτησης πελατών. Αυτές οι μετρήσεις μπορούν να περιλαμβάνουν: Απαντήσεις κλήσεων, Εγκατάλειψη κλήσεων, Μέση διάρκεια χειρισμού και Μέση αναμονή.

Η ανάπτυξη μετρήσεων απόδοσης συνήθως ακολουθεί μια διαδικασία:

- Καθορισμός κρίσιμων διαδικασιών / απαιτήσεων πελατών
- Προσδιορισμός συγκεκριμένων ποσοτικοποιήσιμων αποτελεσμάτων εργασίας
- Καθορισμός στόχων βάσει των οποίων μπορούν να βαθμολογηθούν τα αποτελέσματα.

Μια κριτική των μετρήσεων απόδοσης είναι ότι όταν η αξία των πληροφοριών υπολογίζεται χρησιμοποιώντας μαθηματικές μεθόδους, δείχνει ότι ακόμη και οι επαγγελματίες των μετρήσεων απόδοσης επιλέγουν μέτρα που έχουν μικρή αξία. Αυτό αναφέρεται ως «αντιστροφή μέτρησης». Για παράδειγμα, οι μετρήσεις δείχνουν ότι οι οργανώσεις βρίσκουν άμεσα μετρήσιμες - ακόμη και αν αυτές είναι χαμηλές - και τείνουν να αγνοούν τις μετρήσεις υψηλής αξίας απλώς και μόνο επειδή φαίνεται πιο δύσκολο να μετρηθούν (ανεξάρτητα αν είναι ή όχι).

Για να διορθωθεί η αντιστροφή μέτρησης, άλλες μέθοδοι, όπως η εφαρμοσμένη οικονομική πληροφόρηση, εισάγουν το βήμα της «αξίας της ανάλυσης πληροφοριών» στη διαδικασία έτσι ώστε οι μετρήσεις να επικεντρώνονται σε μέτρα υψηλής αξίας. Οι

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας οργανισμοί στους οποίους έχει εφαρμοστεί αυτό, βρίσκουν ότι ορίζουν εντελώς διαφορετικές μετρήσεις από ότι θα είχαν διαφορετικά και συχνά λιγότερες μετρήσεις. Για τα έργα, η προσπάθεια συλλογής μιας μέτρησης πρέπει να σταθμιστεί έναντι της αξίας της, καθώς τα έργα είναι προσωρινές προσπάθειες που εκτελούνται με πεπερασμένους πόρους.

Υπάρχουν διάφοροι τρόποι με τους οποίους οι οργανισμοί μπορούν να αντιδρούν στα αποτελέσματα. Αυτό μπορεί να είναι η ενεργοποίηση συγκεκριμένης δραστηριότητας σχετικά με τις επιδόσεις (δηλαδή ενός σχεδίου βελτίωσης) ή η χρήση των δεδομένων μόνο για στατιστικές πληροφορίες. Συχνά συνδέονται στενά με τις εξόδους, οι μετρήσεις απόδοσης πρέπει συνήθως να ενθαρρύνουν τη βελτίωση, την αποτελεσματικότητα και τα κατάλληλα επίπεδα ελέγχου (Κύρκος, 2015).

Οι μετρήσεις απόδοσης συχνά συνδέονται με την εταιρική στρατηγική και συχνά προκύπτουν προκειμένου να μετρηθεί η απόδοση σε σχέση με έναν κρίσιμο παράγοντα επιτυχίας.

4.6.2 Συγκριτική αξιολόγηση

Η συγκριτική αξιολόγηση είναι η πρακτική της σύγκρισης των επιχειρηματικών διαδικασιών και των μετρήσεων απόδοσης με τις καλύτερες πρακτικές και τις βέλτιστες πρακτικές άλλων εταιρειών. Οι διαστάσεις που συνήθως μετριοούνται είναι η ποιότητα, ο χρόνος και το κόστος.

Η συγκριτική αξιολόγηση χρησιμοποιείται για τη μέτρηση της απόδοσης χρησιμοποιώντας έναν συγκεκριμένο δείκτη (κόστος ανά μονάδα μέτρησης, παραγωγικότητα ανά μονάδα μέτρησης, χρόνος κύκλου ανά μονάδα μέτρησης ή ελαττώματα ανά μονάδα μέτρησης) με αποτέλεσμα μια μέτρηση απόδοσης που στη συνέχεια συγκρίνεται με άλλες.

Αναφέρεται επίσης ως «συγκριτική αξιολόγηση των βέλτιστων πρακτικών» ή «συγκριτική αξιολόγηση των διαδικασιών», η διαδικασία αυτή χρησιμοποιείται στη διαχείριση, η οποία παρουσιάζει ιδιαίτερα στρατηγική διαχείριση VEMR, στην οποία οι οργανισμοί αξιολογούν διάφορες πτυχές των διαδικασιών τους σε σχέση με τις διαδικασίες των εταιρειών βέλτιστης πρακτικής. Αυτό επιτρέπει στους οργανισμούς να αναπτύσσουν σχέδια σχετικά με τον τρόπο βελτίωσης ή προσαρμογής συγκεκριμένων βέλτιστων πρακτικών, συνήθως με στόχο την αύξηση ορισμένων επιδόσεων. Η συγκριτική αξιολόγηση μπορεί να είναι ένα έκτακτο γεγονός, αλλά αντιμετωπίζεται συχνά ως μια συνεχής διαδικασία στην οποία οι οργανισμοί προσπαθούν συνεχώς να βελτιώνουν τις πρακτικές τους (Κύρκος, 2015).

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

Στη διαχείριση έργων η συγκριτική αξιολόγηση μπορεί επίσης να υποστηρίξει την επιλογή, το σχεδιασμό και την υλοποίηση έργων.

Στη διαδικασία συγκριτικής αξιολόγησης των βέλτιστων πρακτικών, η διοίκηση εντοπίζει τις καλύτερες επιχειρήσεις στον κλάδο τους ή σε μια άλλη βιομηχανία όπου υπάρχουν παρόμοιες διαδικασίες και συγκρίνει τα αποτελέσματα και τις διαδικασίες αυτών που μελετήθηκαν (οι «στόχοι») με τα δικά τους αποτελέσματα και διαδικασίες. Με αυτόν τον τρόπο μαθαίνουν πόσο καλά εκτελούν οι στόχοι και, το σημαντικότερο, τις επιχειρηματικές διαδικασίες που εξηγούν γιατί οι επιχειρήσεις αυτές είναι επιτυχείς. Σύμφωνα με το Εθνικό Συμβούλιο για την Εκτίμηση της Εκπαίδευσης, οι εκτιμήσεις αναφοράς είναι σύντομες αξιολογήσεις που χρησιμοποιούν οι εκπαιδευτικοί σε διάφορες χρονικές στιγμές καθ' όλη τη διάρκεια του σχολικού έτους για την παρακολούθηση της προόδου των σπουδαστών σε κάποια περιοχή του σχολικού προγράμματος σπουδών. Αυτές είναι επίσης γνωστές ως ενδιάμεσες αξιολογήσεις.

Η συγκριτική αξιολόγηση χρησιμοποιείται περισσότερο για τη μέτρηση των επιδόσεων χρησιμοποιώντας έναν συγκεκριμένο δείκτη (κόστος ανά μονάδα μέτρησης, παραγωγικότητα ανά μονάδα μέτρησης, χρόνος κύκλου ανά μονάδα μέτρησης ή ελαττώματα ανά μονάδα μέτρησης) με αποτέλεσμα μια μέτρηση απόδοσης που στη συνέχεια συγκρίνεται με τους υπόλοιπους (Κύρκος, 2015).

4.6.3 Διαχείριση γνώσης

Η «διαχείριση γνώσης», είναι η διαδικασία δημιουργίας, διανομής, χρήσης και διαχείρισης των γνώσεων και πληροφοριών ενός οργανισμού. Αναφέρεται σε μια πολυεπιστημονική προσέγγιση για την επίτευξη των οργανωτικών στόχων με την καλύτερη αξιοποίηση της γνώσης.

Καθορισμένη πειθαρχία από το 1991, η «διαχείριση γνώσης» περιλαμβάνει μαθήματα που διδάσκονται στους τομείς της διοίκησης επιχειρήσεων, των συστημάτων πληροφορικής, της διαχείρισης, της βιβλιοθήκης και των επιστημών της πληροφορίας. Άλλοι τομείς μπορούν να συμβάλουν στην έρευνα του «διαχείριση γνώσης», συμπεριλαμβανομένων των πληροφοριών και των μέσων ενημέρωσης, της πληροφορικής, της δημόσιας υγείας και της δημόσιας τάξης. Αρκετά πανεπιστήμια προσφέρουν πτυχία αφιερωμένα στον τομέα της διαχείρισης γνώσης.

Πολλές μεγάλες εταιρείες, δημόσιοι οργανισμοί και μη κερδοσκοπικοί οργανισμοί διαθέτουν πόρους αφιερωμένους στις εσωτερικές προσπάθειες «διαχείριση γνώσης», συχνά

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

ως μέρος της επιχειρηματικής τους στρατηγικής, των τμημάτων διαχείρισης ανθρωπίνων πόρων. Αρκετές εταιρείες συμβούλων παρέχουν συμβουλές σχετικά με τη «διαχείριση γνώσης» σε αυτούς τους οργανισμούς.

Οι προσπάθειες διαχείρισης της γνώσης επικεντρώνονται συνήθως σε οργανωτικούς στόχους όπως η βελτίωση των επιδόσεων, το ανταγωνιστικό πλεονέκτημα, η καινοτομία, η ανταλλαγή διδαγμάτων, η ολοκλήρωση και η συνεχής βελτίωση του οργανισμού. Αυτές οι προσπάθειες αλληλεπικαλύπτονται με την οργανωτική μάθηση και μπορούν να διακριθούν από αυτό με μεγαλύτερη έμφαση στη διαχείριση της γνώσης ως στρατηγικό πλεονέκτημα και στην ενθάρρυνση της ανταλλαγής γνώσεων. Η «διαχείριση γνώσης» είναι ένας παράγοντας που διευκολύνει την οργανωτική μάθηση (Κύρκος, 2015).

5 Κεφάλαιο: «Επιχειρηματική ευφυΐα σε πραγματικό χρόνο»

5.1 Γενικά στοιχεία

Η Επιχειρηματική Ευφυΐα (Operational intelligence / OI) είναι μια κατηγορία δυναμικών, επιχειρησιακών αναλυτικών στοιχείων σε πραγματικό χρόνο, που προσφέρει ορατότητα και διορατικότητα στα δεδομένα, τις εκδηλώσεις ροής και τις επιχειρηματικές λειτουργίες. Οι λύσεις της OI εκτελούν ερωτήματα κατά των ροών δεδομένων και των συμβάντων για την παροχή αναλυτικών αποτελεσμάτων ως λειτουργικές οδηγίες. Το OI παρέχει στους οργανισμούς τη δυνατότητα να λαμβάνουν αποφάσεις και να ενεργούν άμεσα με αυτές τις αναλυτικές γνώσεις, με χειροκίνητες ή αυτοματοποιημένες ενέργειες (Radenković, Lukić, Despotović-Zrakić, Labus, & Bogdanović, 2018).

5.2 Σκοπός Επιχειρηματικής Ευφυΐας

Σκοπός της Επιχειρηματικής Ευφυΐας είναι να παρακολουθεί τις επιχειρηματικές δραστηριότητες, να εντοπίζει και να ανιχνεύει καταστάσεις που σχετίζονται με αναποτελεσματικότητα, ευκαιρίες και απειλές και να παρέχει επιχειρησιακές λύσεις. Μερικοί ορισμοί ορίζουν την επιχειρησιακή νοημοσύνη ως μια προσέγγιση με βάση την εκδήλωση για την παροχή πληροφοριών που εξουσιοδοτούν τους ανθρώπους να λαμβάνουν καλύτερες αποφάσεις, με βάση πλήρεις και πραγματικές πληροφορίες (García deSoto & Adey, 2016).

Επιπλέον, αυτές οι μετρήσεις χρησιμεύουν ως σημείο εκκίνησης για περαιτέρω ανάλυση (ανάλυση των λεπτομερειών, διεξαγωγή ανάλυσης βασικών αιτίων - σύνδεση των ανωμαλιών με συγκεκριμένες συναλλαγές και επιχειρηματική δραστηριότητα).

Τα εξελιγμένα συστήματα Επιχειρηματικής Ευφυΐας παρέχουν επίσης τη δυνατότητα σύνδεσης των μεταδεδομένων με μετρήσεις, βήματα επεξεργασίας, κανάλια κ.λπ. Με αυτό, γίνεται εύκολη η λήψη σχετικών πληροφοριών, για παράδειγμα (Κτόλ, 2016): *«ανακτήσετε τα στοιχεία επικοινωνίας του ατόμου που διαχειρίζεται την εφαρμογή το οποίο εκτέλεσε το βήμα με την επιχειρηματική συναλλαγή που χρειάστηκε 60% περισσότερο χρόνο από τον κανόνα»* ή *«δείτε την τάση αποδοχής / απόρριψης για τον πελάτη στον οποίο δεν έγινε δεκτή η έγκριση σε αυτή τη συναλλαγή»* ή *«ξεκινήστε την εφαρμογή με την οποία το βήμα της διαδικασίας αλληλεπίδρασε»*.

5.3 Διάφορα στοιχεία Επιχειρηματικής Ευφυΐας

Οι διαφορετικές λειτουργικές λύσεις πληροφοριών μπορούν να χρησιμοποιούν πολλές διαφορετικές τεχνολογίες και να υλοποιούνται με διαφορετικούς τρόπους. Παρακάτω παρατίθενται τα κοινά χαρακτηριστικά μιας επιχειρησιακής λύσης πληροφοριών (Κρόι, 2016):

- Παρακολούθηση σε πραγματικό χρόνο
- Ανίχνευση καταστάσεων πραγματικού χρόνου
- Πίνακες ελέγχου σε πραγματικό χρόνο για διαφορετικούς ρόλους χρηστών
- Συσχέτιση των γεγονότων
- Ειδικοί πίνακες ελέγχου για τη βιομηχανία
- Πολυδιάστατη ανάλυση
- Ανάλυση των βαθύτερων αιτίων
- Χρονική σειρά και ανάλυση τάσεων
- Μεγάλα δεδομένα Analytics: Η επιχειρησιακή ευφυΐα είναι κατάλληλη για την αντιμετώπιση των εγγενών προκλήσεων του Big Data. Επίσης, η επιχειρησιακή νοημοσύνη παρακολουθεί συνεχώς και αναλύει την ποικιλία των μεγάλων πηγών δεδομένων μεγάλης ταχύτητας και μεγάλου όγκου. Συχνά εκτελούνται στη μνήμη οι πλατφόρμες και οι λύσεις της Επιχειρησιακής Ευφυΐας παρουσιάζουν τους αυξητικούς υπολογισμούς και τις αλλαγές, σε πραγματικό χρόνο, στον τελικό χρήστη.

5.4 Τεχνολογικά στοιχεία Επιχειρηματικής Ευφυΐας

Οι λύσεις επιχειρησιακής νοημοσύνης μοιράζονται πολλά χαρακτηριστικά και επομένως πολλοί μοιράζονται επίσης τα στοιχεία της τεχνολογίας. Πρόκειται για μια λίστα με ορισμένα από τα συστατικά στοιχεία της τεχνολογίας που βρίσκονται συνήθως στη διάθεσή τους και τα χαρακτηριστικά τους (Radenković, Lukić, Despotović-Zrakić, Labus, & Bogdanović, 2018):

- **Παρακολούθηση επιχειρηματικής δραστηριότητας** (Business activity monitoring / BAM): είναι λογισμικό που βοηθά στην παρακολούθηση επιχειρηματικών δραστηριοτήτων, βασικών δεικτών απόδοσης, επιχειρησιακών / λειτουργικών εξαιρέσεων και επιχειρηματικών κινδύνων, καθώς οι δραστηριότητες αυτές εφαρμόζονται σε συστήματα υπολογιστών.
- **Σύνθετη επεξεργασία συμβάντων** (CEP) (αναλύθηκε στην ενότητα 3.1).

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

- **Διαχείριση επιχειρησιακών διαδικασιών** (Business process management / BPM): είναι μια πειθαρχία στη διαχείριση επιχειρήσεων, στην οποία οι άνθρωποι χρησιμοποιούν διάφορες μεθόδους για να ανακαλύψουν, να μοντελοποιήσουν, να αναλύσουν, να μετρήσουν, να βελτιώσουν και να αυτοματοποιήσουν τις επιχειρηματικές διαδικασίες.
- Το **πλαίσιο μεταδεδομένων**: θα μοντελοποιεί και θα συνδέει τα γεγονότα με τους πόρους.
- **Βάση δεδομένων** (data warehouse): είναι ένα σύστημα που χρησιμοποιείται για την αναφορά και την ανάλυση δεδομένων και θεωρείται βασικό στοιχείο της επιχειρηματικής ευφυΐας
- **Ανάλυση των βαθύτερων αιτιών** (Root cause analysis / RCA): είναι μια μέθοδος επίλυσης προβλημάτων που χρησιμοποιείται για τον εντοπισμό των βασικών αιτιών βλαβών ή προβλημάτων. Ένας παράγοντας θεωρείται πρωταρχική αιτία αν η απομάκρυνσή του από την αλληλουχία σφάλματος-προβλήματος εμποδίζει την επαναλαμβανόμενη τελική ανεπιθύμητη έκβαση.

Η επιχειρησιακή υπηρεσία πληροφοριών είναι ένα σχετικά νέο τμήμα της αγοράς (σε σύγκριση με τα πιο παλιά τμήματα διαχείρισης επιχειρηματικών πληροφοριών και επιχειρηματικών διαδικασιών). Εκτός από τις εταιρείες που παράγουν αποκλειστικά προϊόντα σε αυτόν τον τομέα, υπάρχουν πολλές εταιρείες σε παρακείμενες περιοχές που παρέχουν λύσεις με ορισμένα συστατικά στοιχεία της BI (Radenković, Lukić, Despotović-Zrakić, Labus, & Bogdanović, 2018).

Η επιχειρησιακή νοημοσύνη ενσωματώνει πληροφορίες, υποστηρίζοντας έγκαιρα τη λήψη ευφύστερων αποφάσεων ώστε να μεγιστοποιηθεί ο αντίκτυπος. Με τη συσχέτιση ποικίλων γεγονότων και δεδομένων τόσο από ροές όσο και από ιστορικά δεδομένα, η επιχειρησιακή ευφυΐα βοηθά τους οργανισμούς να αποκτήσουν σε πραγματικό χρόνο προβολή των πληροφοριών, μέσω των dashboards, την πραγματική εικόνα των επιχειρηματικών επιδόσεων με βάση τις επιχειρηματικές πολιτικές και διαδικασίες. Ακόμα η επιχειρησιακή νοημοσύνη εφαρμόζει τα οφέλη της ανάλυσης σε πραγματικό χρόνο, των προειδοποιήσεων και των ενεργειών σε ένα ευρύ φάσμα περιπτώσεων χρήσης πέρα από την επιχείρηση.

Ένα συγκεκριμένο τμήμα της τεχνολογίας είναι η Αυτόματη αναγνώριση και καταγραφή δεδομένων (Automatic identification and data capture / AIDC) που αντιπροσωπεύεται από γραμμωτούς κώδικες (barcode), η RFID (Radio-frequency identification / η αναγνώριση ραδιοσυχνοτήτων που χρησιμοποιεί ηλεκτρομαγνητικά πεδία

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας για την αυτόματη αναγνώριση και παρακολούθηση των ετικετών που συνδέονται με αντικείμενα) και η αναγνώριση φωνής (το διεπιστημονικό υποπεδίο της υπολογιστικής γλωσσολογίας που αναπτύσσει μεθοδολογίες και τεχνολογίες που επιτρέπουν την αναγνώριση και τη μετάφραση της ομιλούμενης γλώσσας σε κείμενο από τους υπολογιστές). Μια άλλη συγκεκριμένη τεχνολογία είναι η πλατφόρμα OKAPI. Πρόκειται για μια πλατφόρμα λογισμικού επιχειρησιακής αριστείας που χρησιμοποιεί την τεχνητή νοημοσύνη και τη μηχανική μάθηση για να παρέχει στις επιχειρήσεις KPIs SMART. Στη συνέχεια, η πλατφόρμα χρησιμοποιεί την οπτικοποίηση δεδομένων για να παρακολουθεί την πρόοδο του χτύπου των KPIs (García deSoto & Adey, 2016).

5.5 Επιχειρηματική νοημοσύνη σε πραγματικό χρόνο

Η Επιχειρηματική νοημοσύνη σε πραγματικό χρόνο (Real-time business intelligence / RTBI) είναι μια έννοια που περιγράφει τη διαδικασία παροχής επιχειρηματικών πληροφοριών ή πληροφορίες σχετικά με επιχειρηματικές λειτουργίες όπως αυτές συμβαίνουν. Ο πραγματικός χρόνος σημαίνει σχεδόν μηδενική καθυστέρηση και πρόσβαση σε πληροφορίες όποτε απαιτείται (Radenković, Lukić, Despotović-Zrakić, Labus, & Bogdanović, 2018).

Η ταχύτητα των σημερινών συστημάτων επεξεργασίας επέτρεψε την τυπική αποθήκευση δεδομένων να λειτουργεί σε πραγματικό χρόνο. Το αποτέλεσμα είναι επιχειρηματική ευφυΐα σε πραγματικό χρόνο. Οι επιχειρηματικές συναλλαγές όπως συμβαίνουν τροφοδοτούνται σε ένα σύστημα BI σε πραγματικό χρόνο το οποίο διατηρεί την τρέχουσα κατάσταση της επιχείρησης. Το σύστημα RTBI υποστηρίζει όχι μόνο τις κλασικές στρατηγικές λειτουργίες της αποθήκευσης δεδομένων για την παραγωγή πληροφοριών και γνώσεων από παλαιότερες επιχειρηματικές δραστηριότητες, αλλά παρέχει και τακτική υποστήριξη σε πραγματικό χρόνο για την προώθηση ενεργειών επιχειρήσεων που αντιδρούν άμεσα στα γεγονότα όπως αυτά συμβαίνουν. Ως εκ τούτου, αντικαθιστά τόσο την κλασική αποθήκη δεδομένων όσο και τις λειτουργίες ολοκλήρωσης εφαρμογών για επιχειρήσεις (Enterprise application integration / EAI). Τέτοιες επεξεργασίες που βασίζονται σε συμβάντα αποτελούν βασική αρχή επιχειρηματικής ευφυΐας σε πραγματικό χρόνο.

Σε αυτό το πλαίσιο, «σε πραγματικό χρόνο» σημαίνει μια περιοχή από χιλιοστά του δευτερολέπτου έως μερικά δευτερόλεπτα (5 δευτερόλεπτα) μετά την πραγματοποίηση του επιχειρηματικού γεγονότος. Ενώ το παραδοσιακό BI παρουσιάζει ιστορικά δεδομένα για χειρωνακτική ανάλυση, το RTBI συγκρίνει τα τρέχοντα επιχειρηματικά γεγονότα με ιστορικά

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας μοντέλα για την αυτόματη ανίχνευση προβλημάτων ή ευκαιριών. Αυτή η αυτοματοποιημένη δυνατότητα ανάλυσης επιτρέπει την έναρξη διορθωτικών ενεργειών και / ή την προσαρμογή των επιχειρηματικών κανόνων για τη βελτιστοποίηση των επιχειρηματικών διαδικασιών.

Το RTBI είναι μια προσέγγιση στην οποία αναλύονται τα δεδομένα που είναι διαθέσιμα σε λεπτά, είτε απευθείας από επιχειρησιακές πηγές είτε με την παροχή επιχειρηματικών συναλλαγών σε μια αποθήκη δεδομένων πραγματικού χρόνου και στο σύστημα Business Intelligence (Radenković, Lukić, Despotović-Zrakić, Labus, & Bogdanović, 2018).

5.6 Λήψη αποφάσεων Επιχειρηματικής ευφυΐας

Με τις υψηλές προσδοκίες των καταναλωτών στην ανταγωνιστική αγορά, οι αποφάσεις που βασίζονται στα πιο διαθέσιμα δεδομένα σε πραγματικό χρόνο βελτιώνουν τις σχέσεις με τους πελάτες, αυξάνουν τα έσοδα, αυξάνουν την αποδοτικότητα για την αναφορά των επιχειρήσεων και μεγιστοποιούν την επιχειρησιακή αποτελεσματικότητα. Συστήματα πληροφοριών επιχειρησιακών πληροφοριών σε πραγματικό χρόνο παρέχουν κυρίως τις απαραίτητες πληροφορίες για την τακτική εκμετάλλευση των γεγονότων όπως συμβαίνουν.

Σε επίπεδο ανάλυσης, τα συστήματα BI παρέχουν ενοποιημένες πληροφορίες που οι χρήστες μπορούν να επεξεργαστούν λεπτομερώς και να φιλτράρουν, παρέχει επίσης αναλύσεις προβλέψεων και αναλύσεις τάσεων για την ανάπτυξη νέων επιχειρηματικών στοιχείων (βασισμένων στα πρωτογενή δεδομένα) (Σταλίδης & Καρδαράς, 2015).

5.7 Καθυστέρηση Επιχειρηματικής ευφυΐας

Όλα τα συστήματα επιχειρησιακής νοημοσύνης σε πραγματικό χρόνο έχουν κάποια καθυστέρηση αλλά ο στόχος είναι να ελαχιστοποιηθεί ο χρόνος από το επιχειρηματικό γεγονός που συμβαίνει σε μια διορθωτική ενέργεια ή μια ειδοποίηση που ξεκινάει. Ο αναλυτής Richard Hackathorn περιγράφει τρεις τύπους λανθάνοντος χρόνου (Σταλίδης & Καρδαράς, 2015).

- Λανθάνουσα κατάσταση δεδομένων. Πρόκειται για τον χρόνο που απαιτείται για τη συλλογή και αποθήκευση των δεδομένων.
- Λανθάνουσα ανάλυση. Πρόκειται για τον χρόνο που απαιτείται για την ανάλυση των δεδομένων και η μετατροπή τους σε πληροφορίες που μπορούν να ενεργοποιηθούν

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

- Λανθάνουσα ενέργεια. Πρόκειται για τον χρόνο που απαιτείται για να αντιδράσει στις πληροφορίες και να αναληφθεί δράση.

Οι τεχνολογίες επιχειρηματικής ευφυΐας σε πραγματικό χρόνο έχουν σχεδιαστεί για να μειώνουν όσο το δυνατόν περισσότερο τα τρία λανθάνοντα χρονικά περιθώρια, ενώ η παραδοσιακή επιχειρησιακή ευφυΐα επιδιώκει να μειώσει την λανθάνουσα κατάσταση των δεδομένων και δεν αντιμετωπίζει την λανθάνουσα ανάλυση ή την καθυστέρηση της δράσης, καθώς και οι δύο διέπονται από χειρωνακτικές διαδικασίες.

Ορισμένοι σχολιαστές έχουν εισαγάγει την έννοια της σωστής επιχειρησιακής νοημοσύνης, η οποία προτείνει ότι οι πληροφορίες θα πρέπει να παραδίδονται λίγο πριν απαιτηθεί και όχι απαραίτητα σε πραγματικό χρόνο.

5.8 Αρχιτεκτονικές Επιχειρηματικής ευφυΐας

5.8.1 Βάση συμβάντων

Σε πραγματικό χρόνο τα συστήματα Business Intelligence χρησιμοποιούνται σε(Radenković, Lukić, Despotović-Zrakić, Labus, & Bogdanović, 2018):

1. **Σύνθετη επεξεργασία συμβάντων** (Complex event processing / CEP): (αναλύθηκε στην ενότητα 3.1).
2. **Επεξεργασία ροής συμβάντων** (Event stream processing, or ESP): είναι ένα σύνολο τεχνολογιών που σχεδιάστηκαν για να βοηθήσουν στην κατασκευή πληροφοριακών συστημάτων που βασίζονται σε γεγονότα. Οι τεχνολογίες ESP περιλαμβάνουν οπτικοποίηση συμβάντων, βάσεις δεδομένων συμβάντων, μεσαίο λογισμικό που βασίζεται σε γεγονότα και γλώσσες επεξεργασίας συμβάντων ή περίπλοκη επεξεργασία συμβάντων (CEP). Στην πράξη, οι όροι ESP και CEP χρησιμοποιούνται συχνά εναλλακτικά. Το ESP ασχολείται με την επεξεργασία ροών δεδομένων συμβάντων με στόχο τον εντοπισμό του σημαντικού προτύπου μέσα σε αυτά τα ρεύματα, χρησιμοποιώντας τεχνικές όπως ανίχνευση σχέσεων μεταξύ πολλαπλών γεγονότων, συσχέτιση συμβάντων, ιεραρχίες γεγονότων και άλλες πτυχές όπως αιτιότητα, συμμετοχή και χρονική στιγμή.
3. **Mashup** (υβριδική web εφαρμογή): είναι μια ιστοσελίδα ή web εφαρμογή που χρησιμοποιεί περιεχόμενο από περισσότερες από μία πηγές για τη δημιουργία ενός ενιαίου νέα υπηρεσία που εμφανίζονται σε ένα ενιαίο γραφικό περιβάλλον. Για παράδειγμα, ένας χρήστης θα μπορούσε να συνδυάσει τις διευθύνσεις και τις

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

φωτογραφίες των υποκαταστημάτων της βιβλιοθήκης με έναν χάρτη της Google για να δημιουργήσει ένα mashup χάρτη. Ο όρος συνεπάγεται εύκολη και γρήγορη ενσωμάτωση, συχνά χρησιμοποιώντας ανοικτές διεπαφές προγραμματισμού εφαρμογών (open API) και πηγές δεδομένων για την παραγωγή εμπλουτισμένων αποτελεσμάτων που δεν ήταν απαραίτητα ο αρχικός λόγος για την παραγωγή των δεδομένων ακατέργαστης πηγής.

Οι παραπάνω τεχνικές, για να μπορέσουν να αναλύσουν γεγονότα / στοιχεία / δεδομένα, χωρίς να τα μετατρέψουν πρώτα τα αποθηκεύουν σε μια βάση δεδομένων. Αυτές οι τεχνικές βάσεων δεδομένων στη μνήμη έχουν το πλεονέκτημα ότι μπορούν να παρακολουθούνται υψηλά ποσοστά συμβάντων και δεδομένου ότι τα δεδομένα δεν χρειάζεται να γραφούν σε βάσεις δεδομένων, η λανθάνουσα κατάσταση των δεδομένων μπορεί να μειωθεί σε χιλιοστά του δευτερολέπτου.

5.8.2 Αποθήκη δεδομένων

Μια εναλλακτική προσέγγιση στις αρχιτεκτονικές που βασίζονται σε συμβάντα είναι η αύξηση του κύκλου ανανέωσης μιας υπάρχουσας αποθήκης δεδομένων για την ταχύτερη ενημέρωση των δεδομένων. Αυτά τα συστήματα αποθήκευσης δεδομένων σε πραγματικό χρόνο μπορούν να επιτύχουν σχεδόν ενημέρωση σε πραγματικό χρόνο των δεδομένων, όπου η λανθάνουσα κατάσταση των δεδομένων κυμαίνεται συνήθως από λεπτά έως ώρες. Η ανάλυση των δεδομένων εξακολουθεί να είναι συνήθως χειροκίνητη, οπότε η συνολική λανθάνουσα κατάσταση διαφέρει σημαντικά από τις αρχιτεκτονικές προσεγγίσεις που καθοδηγούνται από γεγονότα (Radenković, Lukić, Despotović-Zrakić, Labus, & Bogdanović, 2018).

5.8.3 Τεχνολογία χωρίς διακομιστές

Η τελευταία εναλλακτική καινοτομία σε αρχιτεκτονικές αποθήκης δεδομένων «πραγματικού χρόνου» είναι η τεχνολογία MSSO (Multiple Source Simple Output / απλή έξοδος πολλαπλών πηγών) η οποία καταργεί την ανάγκη για αποθήκη δεδομένων και ενδιάμεσους διακομιστές καθότι είναι σε θέση να έχει πρόσβαση ζωντανά δεδομένα απευθείας από την πηγή (ακόμη και από πολλαπλές, διαφορετικές πηγές). Επειδή τα ζωντανά δεδομένα προσπελάζονται απευθείας με μέσα χωρίς διακομιστή, παρέχει τη δυνατότητα για δεδομένα με πραγματική ώρα μηδενικής καθυστέρησης με την πιο πιστή έννοια (Cook, 2014).

5.8.4 Ενημέρωση σχετικά με τη διαδικασία

Αυτό μερικές φορές θεωρείται ως ένα υποσύνολο επιχειρησιακής νοημοσύνης και αναγνωρίζεται επίσης με την παρακολούθηση της επιχειρηματικής δραστηριότητας. Επιτρέπει την παρακολούθηση ολόκληρων διεργασιών (συναλλαγών, βημάτων), μετρήσεων (λανθάνουσα κατάσταση, αναλύσεων ολοκλήρωσης / αποτυχίας κλπ). Σε σύγκριση με τα αποθηκευμένα ιστορικά δεδομένα και την τάση σε πραγματικό χρόνο. Οι προηγμένες εφαρμογές επιτρέπουν την ανίχνευση κατώτατης τιμής, την προειδοποίηση και την παροχή ανατροφοδότησης στα ίδια τα συστήματα εκτέλεσης της διαδικασίας, οπότε «κλείνει το βρόχο» (Cook, 2014).

5.9 Τεχνολογίες που υποστηρίζουν την ανάλυση σε πραγματικό χρόνο

Οι τεχνολογίες που μπορούν να υποστηριχθούν για την παροχή επιχειρηματικής ευφυΐας σε πραγματικό χρόνο είναι η οπτικοποίηση δεδομένων, η ομοσπονδία δεδομένων, η ενοποίηση των επιχειρηματικών πληροφοριών, η ενοποίηση των εφαρμογών των επιχειρήσεων και η αρχιτεκτονική προσανατολισμένη στις υπηρεσίες. Τα σύνθετα εργαλεία επεξεργασίας συμβάντων μπορούν να χρησιμοποιηθούν για την ανάλυση ροών δεδομένων σε πραγματικό χρόνο και είτε να ενεργοποιήσουν αυτοματοποιημένες ενέργειες είτε να προειδοποιήσουν τους εργαζόμενους για τα πρότυπα και τις τάσεις (Radenković, Lukić, Despotović-Zrakić, Labus, & Bogdanović, 2018).

5.9.1 Συσκευή αποθήκης δεδομένων

Η συσκευή αποθήκης δεδομένων είναι ένας συνδυασμός υλικού και λογισμικού το οποίο σχεδιάστηκε αποκλειστικά για αναλυτική επεξεργασία. Στην υλοποίηση της αποθήκης δεδομένων, οι εργασίες που αφορούν τον συντονισμό, την προσθήκη ή την επεξεργασία της δομής γύρω από τα δεδομένα, τη φυγή δεδομένων από άλλες βάσεις δεδομένων, τη σύζευξη δεδομένων πραγματοποιούνται από την DBA (Database administrator / διαχειριστής βάσεων δεδομένων). Ένα άλλο καθήκον για το DBA ήταν να καταστήσει την βάση δεδομένων καλή για τις μεγάλες ομάδες χρηστών. Ενώ με συσκευές αποθήκευσης δεδομένων, είναι ο υπεύθυνος του πωλητή για το φυσικό σχεδιασμό και τον συντονισμό του λογισμικού σύμφωνα με τις απαιτήσεις υλικού. Το πακέτο συσκευών αποθήκης δεδομένων διαθέτει δικό του λειτουργικό σύστημα, αποθήκευση, DBMS (συστήματα διαχείρισης βάσεων δεδομένων), λογισμικό και απαιτούμενο υλικό. Τέλος, οι απαιτούμενες συσκευές αποθήκης δεδομένων

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας μπορούν εύκολα να ενσωματωθούν με άλλα εργαλεία (Radenković, Lukić, Despotović-Zrakić, Labus, & Bogdanović, 2018).

5.9.2 Κινητή τεχνολογία

Υπάρχουν πολύ περιορισμένοι προμηθευτές για την παροχή επιχειρηματικών πληροφοριών κινητής τηλεφωνίας. Το MBI (Mobile Business Intelligence) είναι ενσωματωμένο στην υπάρχουσα αρχιτεκτονική BI. Το MBI είναι ένα πακέτο που χρησιμοποιεί τις υπάρχουσες εφαρμογές BI, ώστε οι χρήστες να μπορούν να χρησιμοποιούν το κινητό τους τηλέφωνο και να λαμβάνουν τεκμηριωμένη απόφαση σε πραγματικό χρόνο (Past of the Future, 2018).

5.10 Περιοχές εφαρμογής Επιχειρηματικής ευφυΐας

Οι περιοχές εφαρμογής Επιχειρηματικής ευφυΐας είναι οι εξής (Radenković, Lukić, Despotović-Zrakić, Labus, & Bogdanović, 2018):

- **Αλγοριθμική συναλλαγή**

Αλγοριθμική συναλλαγή είναι μια μέθοδος εκτέλεσης μιας μεγάλης παραγγελίας χρησιμοποιώντας αυτοματοποιημένες προ-προγραμματισμένες οδηγίες συναλλαγών που αφορούν μεταβλητές όπως ο χρόνος, η τιμή και ο όγκος για την αποστολή μικρών τεμαχίων της παραγγελίας στην αγορά με την πάροδο του χρόνου. Αυτά αναπτύχθηκαν έτσι ώστε οι έμποροι να μην χρειάζεται να παρακολουθούν συνεχώς ένα απόθεμα και να αποστέλλουν επανειλημμένα αυτές τις φέτες με το χέρι. Τα δημοφιλή «algorithms» περιλαμβάνουν Pegged, VWAP, TWAP, έλλειψη υλοποίησης, στόχος Close. Στον εικοστό πρώτο αιώνα, το αλγοριθμικό εμπόριο έχει ελκύσει τόσο τους λιανικούς όσο και τους θεσμικούς εμπόρους. Η αλγοριθμική διαπραγμάτευση δεν είναι μια προσπάθεια να γίνει ένα εμπορικό κέρδος. Είναι απλώς ένας τρόπος ελαχιστοποίησης του κόστους, του αντίκτυπου στην αγορά και του κινδύνου κατά την εκτέλεση μιας παραγγελίας. Χρησιμοποιείται ευρέως από επενδυτικές τράπεζες, συνταξιοδοτικά ταμεία, αμοιβαία κεφάλαια (Radenković, Lukić, Despotović-Zrakić, Labus, & Bogdanović, 2018).

- **Ανίχνευση απάτης**

Σύμφωνα με το νόμο, η απάτη είναι σκόπιμη εξαπάτηση για την εξασφάλιση αθέμιτου ή παράνομου κέρδους ή για να στερηθεί από ένα θύμα νομικό δικαίωμα. Η ίδια η απάτη μπορεί να είναι ένα πολιτικό λάθος, ένα ποινικό, ή να μην μπορεί να προκαλέσει απώλεια

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας χρημάτων, περιουσίας ή νόμιμου δικαιώματος, αλλά εξακολουθεί να αποτελεί στοιχείο άλλου αστικού ή ποινικού δικαίου. Ο σκοπός της απάτης μπορεί να είναι η νομισματική κερδοφορία ή άλλες παροχές, όπως η απόκτηση διαβατηρίου ή ταξιδιωτικού εγγράφου, άδειας οδήγησης ή η χορήγηση υποθήκης μέσω ψευδών δηλώσεων (Radenković, Lukić, Despotović-Zrakić, Labus, & Bogdanović, 2018).

- **Διαχείριση σχέσεων πελατών**

Η διαχείριση σχέσεων πελατών (CRM) είναι μια προσέγγιση για τη διαχείριση της αλληλεπίδρασης μιας εταιρείας με τους τρέχοντες και δυνητικούς πελάτες. Χρησιμοποιεί ανάλυση δεδομένων σχετικά με την ιστορία των πελατών με μια εταιρεία για τη βελτίωση των επιχειρηματικών σχέσεων με τους πελάτες, με ιδιαίτερη έμφαση στη διατήρηση των πελατών και την τελική αύξηση της αύξησης των πωλήσεων (Radenković, Lukić, Despotović-Zrakić, Labus, & Bogdanović, 2018).

- **Διαχείριση απόδοσης**

Διαχείριση απόδοσης είναι μια μεταβλητή στρατηγική τιμολόγησης, βασισμένη στην κατανόηση, την πρόβλεψη και την επίδραση στη συμπεριφορά των καταναλωτών, προκειμένου να μεγιστοποιηθούν τα έσοδα ή τα κέρδη από ένα σταθερό, χρονικά περιορισμένο πόρο. Ως ειδικός κλάδος διαχείρισης εσόδων που επικεντρώνεται στην απογραφή, η διαχείριση της απόδοσης περιλαμβάνει στρατηγικό έλεγχο της απογραφής για να πουλήσει το σωστό προϊόν στον κατάλληλο πελάτη τη σωστή στιγμή για τη σωστή τιμή. Αυτή η διαδικασία μπορεί να οδηγήσει σε διακρίσεις ως προς τις τιμές, στις οποίες οι πελάτες που καταναλώνουν πανομοιότυπα προϊόντα ή υπηρεσίες χρεώνονται διαφορετικές τιμές. Η διαχείριση της απόδοσης είναι μια μεγάλη γενιά εσόδων για αρκετές μεγάλες βιομηχανίες. Ο Robert Crandall, πρώην πρόεδρος και διευθύνων σύμβουλος της American Airlines, έδωσε το όνομά του στη διαχείριση των αποδόσεων και την αποκαλούσε «*την πιο σημαντική τεχνική εξέλιξη στη διαχείριση των μεταφορών από τότε που ίσχυσε η απορρύθμιση*» (Radenković, Lukić, Despotović-Zrakić, Labus, & Bogdanović, 2018).

- **Επικύρωση δεδομένων**

Στην επιστήμη των υπολογιστών, η επικύρωση των δεδομένων είναι η διαδικασία εξασφάλισης ότι τα δεδομένα έχουν υποβληθεί σε καθαρισμό δεδομένων για να εξασφαλίσουν ότι έχουν ποιότητα δεδομένων, δηλαδή ότι είναι σωστά και χρήσιμα. Χρησιμοποιεί ρουτίνες, συχνά αποκαλούμενες «κανόνες επικύρωσης», «περιορισμούς επικύρωσης» ή «έλεγχοι ρουτίνας», που ελέγχουν την ορθότητα, τη σημασία

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας και την ασφάλεια των δεδομένων που εισάγονται στο σύστημα. Οι κανόνες μπορούν να εφαρμοστούν μέσω των αυτοματοποιημένων εγκαταστάσεων ενός λεξικού δεδομένων ή με τη συμπερίληψη ρητής λογικής επικύρωσης του προγράμματος εφαρμογής (Radenković, Lukić, Despotović-Zrakić, Labus, & Bogdanović, 2018).

- **Επιχειρησιακή Ευφυΐα**

Η Επιχειρησιακή Ευφυΐα (BI) είναι μια κατηγορία δυναμικών, επιχειρησιακών αναλυτικών στοιχείων σε πραγματικό χρόνο, που προσφέρει ορατότητα και διορατικότητα στα δεδομένα, τις εκδηλώσεις ροής και τις επιχειρηματικές λειτουργίες. Οι λύσεις της BI εκτελούν ερωτήματα κατά των ροών δεδομένων και των συμβάντων για την παροχή αναλυτικών αποτελεσμάτων ως λειτουργικές οδηγίες. Η BI παρέχει στους οργανισμούς τη δυνατότητα να λαμβάνουν αποφάσεις και να ενεργούν άμεσα με αυτές τις αναλυτικές γνώσεις, με χειροκίνητες ή αυτοματοποιημένες ενέργειες (Radenković, Lukić, Despotović-Zrakić, Labus, & Bogdanović, 2018).

- **Βελτιστοποίηση της αλυσίδας εφοδιασμού**

Η βελτιστοποίηση της αλυσίδας εφοδιασμού είναι η εφαρμογή διαδικασιών και εργαλείων για τη διασφάλιση της βέλτιστης λειτουργίας μιας αλυσίδας εφοδιασμού παραγωγής και διανομής. Αυτό περιλαμβάνει τη βέλτιστη τοποθέτηση αποθεμάτων στην αλυσίδα εφοδιασμού, ελαχιστοποιώντας το κόστος λειτουργίας. Αυτό συχνά συνεπάγεται την εφαρμογή μαθηματικών τεχνικών μοντελοποίησης χρησιμοποιώντας λογισμικό υπολογιστή (Radenković, Lukić, Despotović-Zrakić, Labus, & Bogdanović, 2018).

- **Αισθητήρας δικτύου**

Το ασύρματο δίκτυο αισθητήρων (Wireless sensor network / WSN) αναφέρεται σε μια ομάδα χωρικά διασκορπισμένων και ειδικών αισθητήρων για την παρακολούθηση και καταγραφή των φυσικών συνθηκών του περιβάλλοντος και την οργάνωση των συλλεγόμενων δεδομένων σε κεντρική θέση. Τα WSNs μετρούν τις περιβαλλοντικές συνθήκες όπως η θερμοκρασία, ο ήχος, τα επίπεδα μόλυνσης, η υγρασία, ο άνεμος κ.ά. (Radenković, Lukić, Despotović-Zrakić, Labus, & Bogdanović, 2018).

6 Κεφάλαιο: «Εργαλεία οπτικοποίησης δεδομένων»

6.1 Γενικά στοιχεία

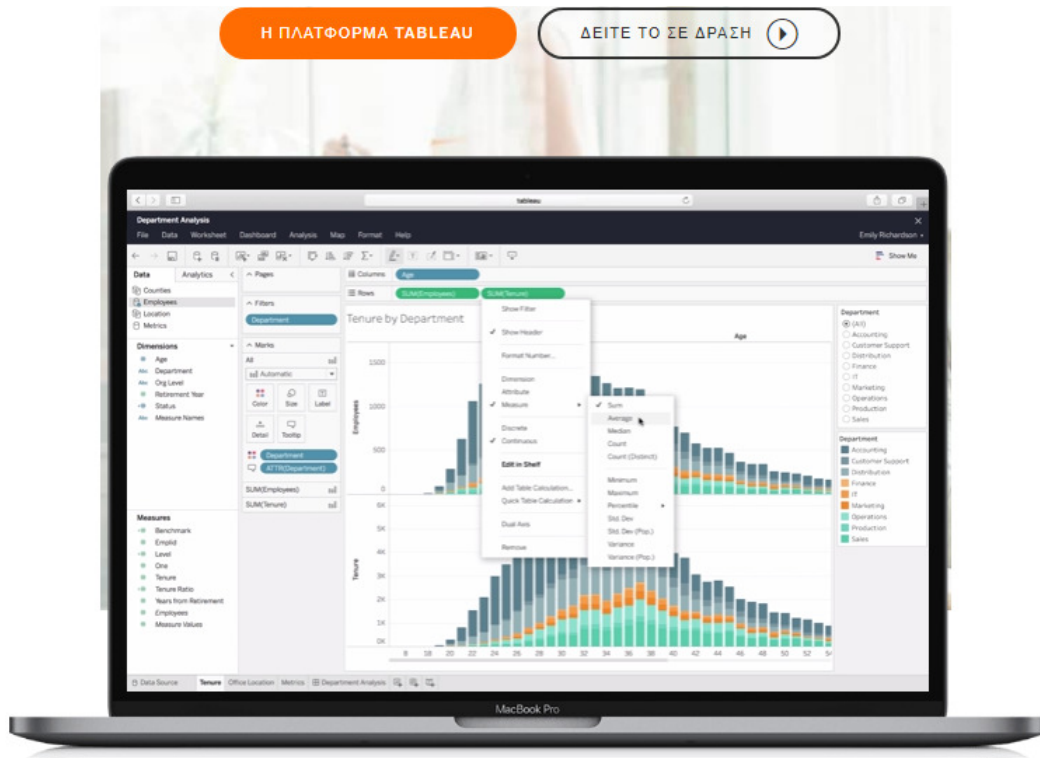
Η απεικόνιση δεδομένων αφορά τον τρόπο παρουσίασης των δεδομένων στους κατάλληλους ανθρώπους, την κατάλληλη στιγμή προκειμένου να μπορέσουν να αποκτήσουν τις καλύτερες γνώσεις.

Η οπτικοποίηση δεδομένων μπορεί να αλλάξει όχι μόνο τον τρόπο με τον οποίο γίνονται αντιληπτά τα δεδομένα αλλά το πόσο γρήγορα και αποτελεσματικά μπορεί να λαμβάνονται αποφάσεις.

Ευτυχώς οι λύσεις απεικόνισης εξελίσσονται τόσο γρήγορα όσο και η υπόλοιπη τεχνολογική πρόοδος. Τα διαγράμματα, τα βίντεο, τα πληροφοριακά στοιχεία ακόμη και οι παρουσιάσεις της εικονικής πραγματικότητας και της επαυξημένης πραγματικότητας (VR & AR) προσφέρουν όλο και πιο συναρπαστικά και διαισθητικά κανάλια επικοινωνίας.

6.2 Tableau

Το Tableau θεωρείται συχνά ως ο μεγάλος κύριος του λογισμικού απεικόνισης δεδομένων και για καλό λόγο. Το Tableau διαθέτει μια πολύ μεγάλη πελατειακή βάση 57.000+ λογαριασμών σε πολλές βιομηχανίες λόγω της απλότητας χρήσης και της ικανότητάς του να παράγει διαδραστικές απεικονίσεις πολύ πιο πέρα από αυτές που παρέχονται από γενικές λύσεις BI. Είναι ιδιαίτερα κατάλληλο για το χειρισμό των τεράστιων και πολύ ταχέως μεταβαλλόμενων συνόλων δεδομένων που χρησιμοποιούνται στις λειτουργίες Big Data, συμπεριλαμβανομένων εφαρμογών τεχνητής νοημοσύνης και μηχανικής μάθησης, χάρη στην ενσωμάτωση με μεγάλο αριθμό προηγμένων λύσεων βάσεων δεδομένων, όπως Hadoop, Amazon AWS, My SQL, SAP και Teradata. Εκτεταμένες έρευνες και δοκιμές επέτρεψαν στο Tableau να δημιουργήσει γραφικά και απεικονίσεις όσο το δυνατόν πιο αποτελεσματικά και να τα καταστήσει εύκολα για τους ανθρώπους να τα καταλάβουν (Tableau, 2019).

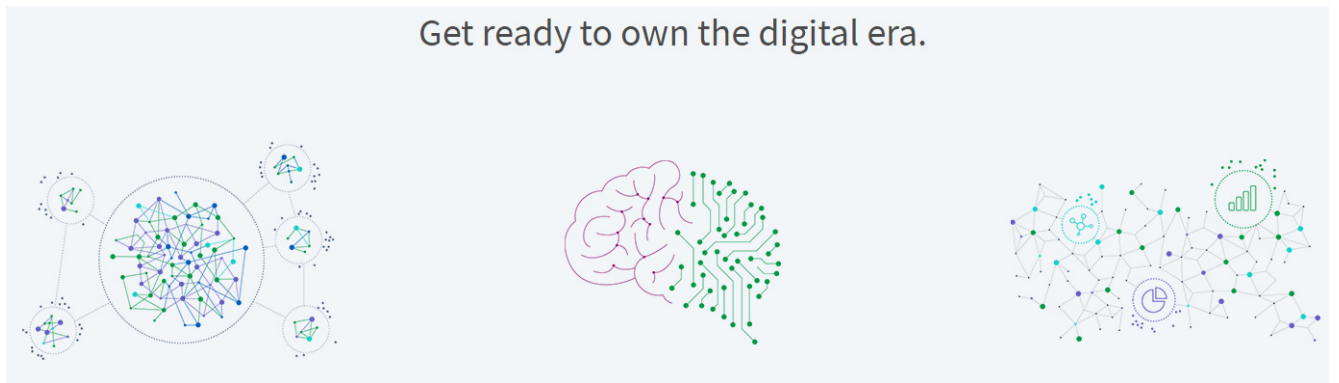


Σχήμα 6.1: Αρχική σελίδα από το λογισμικό Tableau.

Πηγή: (Tableau, 2019).

6.3 Qlikview

Το Qlik με το εργαλείο Qlikview είναι ένα άλλο σημαντικό εργαλείο σε αυτόν τον χώρο και ο μεγαλύτερος ανταγωνιστής του Tableau. Ο πωλητής διαθέτει περισσότερους από 40.000 λογαριασμούς πελατών σε περισσότερες από 100 χώρες, ενώ εκείνοι που το χρησιμοποιούν συχνά επικαλούνται την εξαιρετικά προσαρμόσιμη ρύθμισή του και την ευρεία γκάμα χαρακτηριστικών ως βασικό πλεονέκτημα. Αυτό όμως μπορεί να σημαίνει ότι χρειάζεται περισσότερος χρόνος για να αντιμετωπιστεί και να αξιοποιηθεί πλήρως. Εκτός από τις δυνατότητες απεικόνισης δεδομένων, το Qlikview προσφέρει δυνατότητες επιχειρησιακής ευφυΐας, αναλύσεων και επιχειρηματικών αναφορών και αρέσει ιδιαίτερα καθώς είναι φιλική προς το χρήστη και καθαρή διεπαφή. Το Qlikview χρησιμοποιείται συνήθως μαζί με το πακέτο του QlikSense, το οποίο χειρίζεται την εξερεύνηση και την ανακάλυψη δεδομένων (Qlik, 2019).



Σχήμα 6.2: Αρχική σελίδα από το λογισμικό Qlikview.

Πηγή: (Qlik, 2019).

6.4 FusionCharts

Πρόκειται για ένα πολύ ευρέως χρησιμοποιούμενο εργαλείο, βασισμένο στο JavaScript πακέτο χαρτογράφησης και απεικόνισης που έχει καθιερωθεί ως ένα από τους ηγέτες στην αγορά που καταβάλλεται. Μπορεί να παράγει 90 διαφορετικούς τύπους γραφημάτων και να ενσωματώνεται με μεγάλο αριθμό πλατφορμών και πλαισίων που προσφέρουν μεγάλη ευελιξία. Ένα χαρακτηριστικό που βοήθησε το FusionCharts να γίνει πολύ δημοφιλές είναι ότι αντί να χρειάζεται να ξεκινήσει κάθε νέα απεικόνιση από την αρχή, οι χρήστες μπορούν να επιλέξουν από μια σειρά πρότυπων υποδειγμάτων, απλά συνδέοντας τις δικές τους πηγές δεδομένων, όπως απαιτείται (fusioncharts, 2019).



Σχήμα 6.3: Αρχική σελίδα από το λογισμικό fusioncharts.

Πηγή: (fusioncharts, 2019).

6.5 Highcharts

Όπως και το FusionCharts, έτσι και αυτό το εργαλείο απαιτεί επίσης άδεια για εμπορική χρήση, αν και μπορεί να χρησιμοποιηθεί ελεύθερα για μη εμπορική ή για προσωπική χρήση. Ο ιστότοπός του ισχυρίζεται ότι χρησιμοποιείται από 72 από τις 100 μεγαλύτερες εταιρείες παγκοσμίως και συχνά επιλέγεται όταν πρέπει να αναπτυχθεί μια γρήγορη και ευέλικτη λύση, με ελάχιστη ανάγκη ειδικής κατάρτισης για την απεικόνιση δεδομένων προτού μπορέσει να τεθεί σε λειτουργία. Ένα κλειδί για την επιτυχία του ήταν η επικέντρωσή του στην υποστήριξη πολλαπλών προγραμμάτων περιήγησης, που σημαίνει ότι ο καθένας μπορεί να δει και να εκτελέσει τις διαδραστικές απεικονίσεις, κάτι που δεν ισχύει πάντα με τις νεότερες πλατφόρμες (Highcharts, 2019).

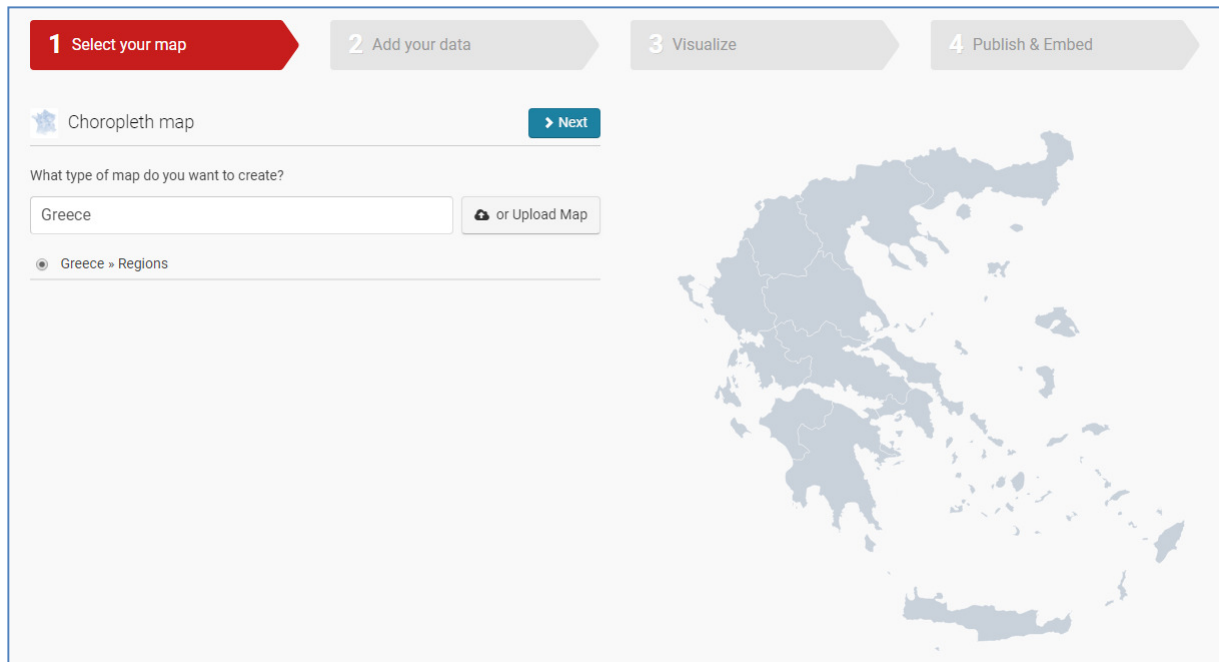


Σχήμα 6.4: Παράδειγμα δημιουργίας διαδραστικών διαγραμμάτων διαφόρων έργων στο διαδίκτυο με το λογισμικό Highcharts.

Πηγή: (Highcharts, 2019).

6.6 Datawrapper

Το Datawrapper γίνεται ολοένα και περισσότερο δημοφιλής επιλογή, ιδίως μεταξύ των οργανώσεων που το χρησιμοποιούν συχνά για να δημιουργούν διαγράμματα και να παρουσιάζουν στατιστικά στοιχεία. Έχει ένα απλό, σαφές interface που καθιστά πολύ εύκολο να ανεβούν δεδομένα csv και να δημιουργηθούν απλά γραφήματα, καθώς και χάρτες, που μπορούν γρήγορα να ενσωματωθούν στις αναφορές.

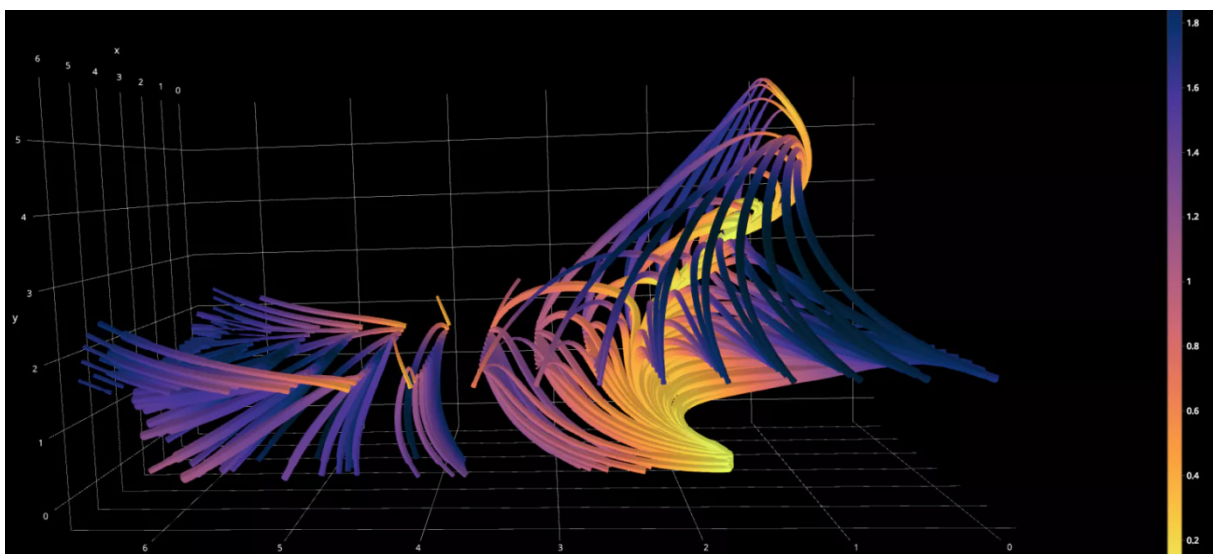


Σχήμα 6.5: Παράδειγμα δημιουργίας διαδραστικών χαρτών διαφόρων δεδομένων στο διαδίκτυο με το λογισμικό Datawrapper.

Πηγή: (Datawrapper, 2019)

6.7 Plotly

Το Plotly επιτρέπει πιο πολύπλοκες και εξελιγμένες οπτικοποιήσεις, χάρη στην ενσωμάτωσή του σε γλώσσες προγραμματισμού προσανατολισμένες σε αναλύσεις όπως οι Python, R και Matlab. Είναι χτισμένο πάνω από τις βιβλιοθήκες οπτικοποίησης d3.js ανοιχτού κώδικα για JavaScript, αλλά αυτό το εμπορικό πακέτο (με μια δωρεάν μη εμπορική άδεια που είναι διαθέσιμη) προσθέτει επίπεδα ευχρηστίας και υποστήριξης καθώς και ενσωματωμένη υποστήριξη για API όπως Salesforce (Plotly, 2019).

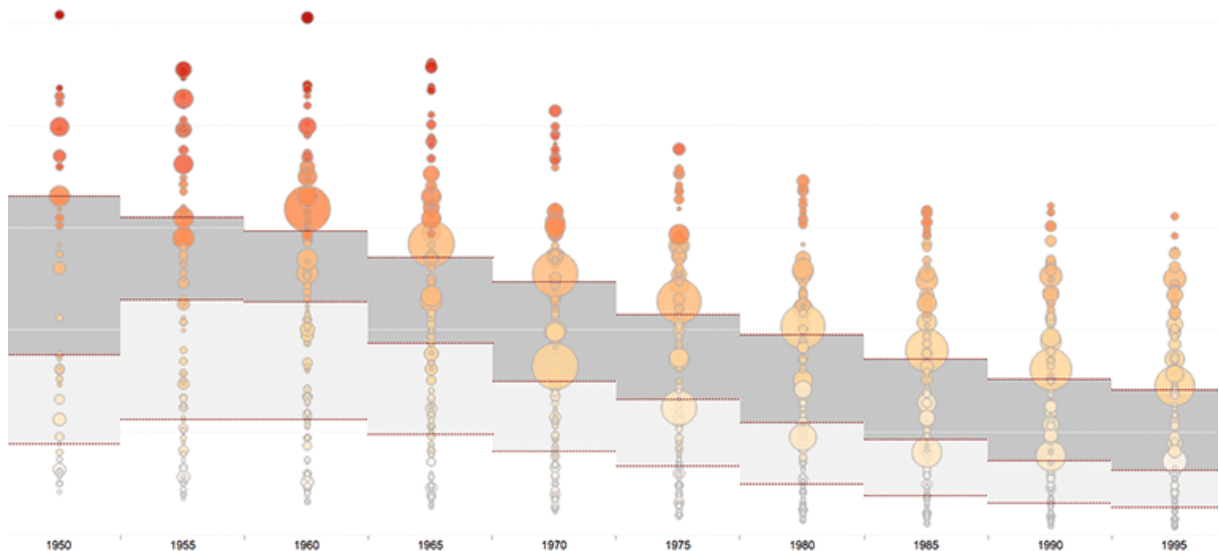


6.8 Sisense

Το Sisense παρέχει μια πλατφόρμα ανάλυσης πλήρους στοίβας, αλλά οι δυνατότητες απεικόνισης του παρέχουν μια απλή στη χρήση εφαρμογή «drag and drop», η οποία επιτρέπει τη δημιουργία γραφικών και πιο σύνθετων γραφικών, καθώς και διαδραστικών απεικονίσεων, με ελάχιστη ταλαιπωρία. Επιτρέπει τη συγκέντρωση πολλαπλών πηγών δεδομένων σε ένα εύκολα προσβάσιμο χώρο αποθήκευσης όπου μπορεί να αναζητηθεί στιγμιαία μέσω των πινάκων ελέγχου, ακόμη και σε σύνολα Big Data. Τα dashboards μπορούν στη συνέχεια να μοιράζονται σε οργανισμούς, εξασφαλίζοντας ότι ακόμη και μη τεχνικά καταρτισμένο προσωπικό μπορεί να βρει τις απαντήσεις που χρειάζονται για τα προβλήματά τους (Sisense, 2019).

6.9 Tableau Public

Αυτό το εργαλείο είναι στην κορυφή γιατί είναι ουσιαστικά η ίδια πλατφόρμα με τη νικήτρια επιλογή. Πρόκειται για την πλήρη έκδοση του Tableau που είναι διαθέσιμη για δωρεάν download, με μία μόνο προειδοποίηση. Τα πάντα που δημιουργούνται μαζί του είναι δημόσια, πράγμα που σημαίνει ότι γίνεται αυτόματα διαθέσιμο στο διαδίκτυο μέσω της γκαλερί απεικόνισης του Tableau (Tableau, 2019).



Εικόνα 6.1: Παράδειγμα αναπαράστασης τακτικών στοιχείων για τη δημόσια υγεία, για τη βελτίωση των ποσοστών θνησιμότητας σε όλο τον κόσμο, με το Tableau Public.

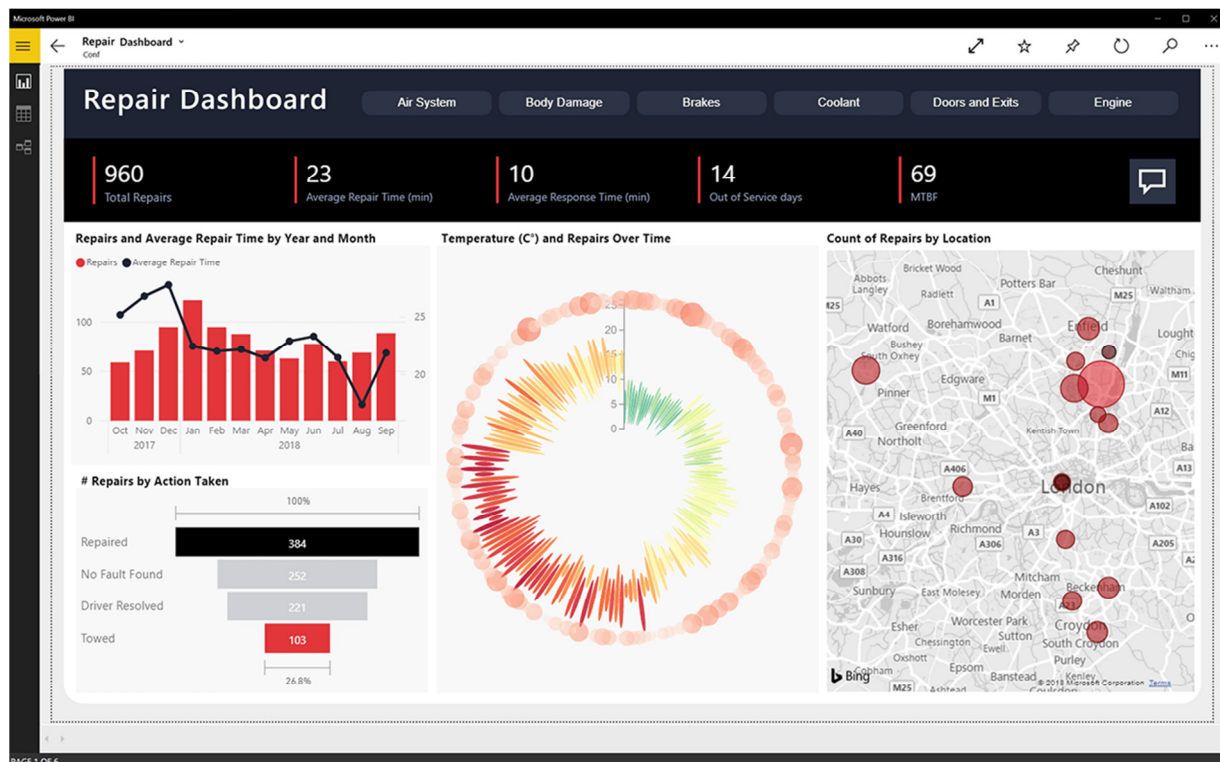
Πηγή:(Tableau, 2019).

6.10 Microsoft Power BI

Πρόκειται για το τελευταίο ριζοκίνδυνο εργαλείο διότι μπορεί να «κατεβεί» δωρεάν. Και επίσης, όπως το Tableau, η Microsoft έχει μια γκαλερί απεικόνισης που μπορεί να προσεγγιστεί τόσο από τους χρήστες της Power BI όσο και από τους ανθρώπους που απλώς αναζητούν ελεύθερα διαθέσιμες απεικονίσεις.

Το Power BI είναι μια υπηρεσία επιχειρηματικών αναλύσεων που παρέχει πληροφορίες για γρήγορες και ενημερωμένες αποφάσεις. Δηλαδή (Microsoft, 2019):

- Μετατρέπει τα δεδομένα σε εκπληκτικά οπτικά και μοιράζονται με συναδέλφους σε οποιαδήποτε συσκευή.
- Εξερευνά οπτικά και αναλύει τα δεδομένα σε εγκαταστάσεις και στο σύννεφο - όλα σε μια προβολή.
- Συνεργάζεται και μοιράζει προσαρμοσμένους πίνακες εργαλείων και διαδραστικές αναφορές.
- Βαθμολογεί τον οργανισμό με ενσωματωμένη διακυβέρνηση και ασφάλεια.



Εικόνα 6.2: Παράδειγμα αναπαράστασης με το λογισμικό Microsoft Power BI.

Πηγή: (Microsoft, 2019).

6.11 Σύγκριση Εργαλείων οπτικοποίησης δεδομένων

Στον παρακάτω πίνακα 6.1 πραγματοποιείται σύγκριση των Εργαλείων οπτικοποίησης δεδομένων, καταγράφοντας τα χαρακτηριστικά τους που τα κάνουν πιο ανταγωνιστικά.

Πίνακας 6.1: Σύγκριση Εργαλείων οπτικοποίησης δεδομένων.

Εργαλεία οπτικοποίησης δεδομένων	Χαρακτηριστικά
Tableau	Απλότητα χρήσης, ενσωμάτωση: Hadoop, Amazon AWS, My SQL, SAP και Teradata
Qlik	Δυνατότητες επιχειρησιακής ευφυΐας, εξαιρετικά προσαρμόσιμη ρύθμιση
FusionCharts	Πακέτο χαρτογράφησης JavaScript, πρότυπα υποδειγμάτα
Highcharts	Ελάχιστη ανάγκη ειδικής κατάρτισης, πολλαπλά προγράμματα περιήγησης
Datawrapper	Απλό interface
Plotly	Επιτρέπει πολύπλοκες και εξελιγμένες οπτικοποιήσεις, βιβλιοθήκες οπτικοποίησης d3.js ανοιχτού κώδικα για JavaScript
Sisense	Διαδραστικές απεικονίσεις, εύκολα προσβάσιμο χώρο αποθήκευσης
Tableau Public	Δωρεάν download
Microsoft Power BI	Μετατρέπει τα δεδομένα σε εκπληκτικά οπτικά, Βαθμολογεί τον οργανισμό με ενσωματωμένη διακυβέρνηση και ασφάλεια

7 Κεφάλαιο: «Συμπεράσματα»

«Η ανάλυση δεδομένων με το πάτημα ενός κουμπιού».

Αυτό ακριβώς θέλουν οι ελεγκτές αλλά και οι έμποροι (επιχειρηματικοί χρήστες). Τα μεγάλα δεδομένα να μπορούν να αξιολογηθούν με το πάτημα ενός κουμπιού. Οι περισσότερες εταιρείες απέχουν ακόμη πολύ από αυτό. Τα παραπάνω παραδείγματα εργαλείων ανάλυσης και οπτικοποίησης δεδομένων που παρουσιάστηκαν δείχνουν πώς μπορεί να «κλείσει» το χάσμα μεταξύ «δεδομένων» και «επιχειρηματικών χρηστών».

Η ανάλυση δεδομένων είναι μια διαδικασία για την απόκτηση ανεπεξέργαστων δεδομένων και τη μετατροπή τους σε πληροφορίες χρήσιμες για τη λήψη αποφάσεων από τους χρήστες. Τα δεδομένα συλλέγονται και αναλύονται για να απαντούν σε ερωτήσεις, να δοκιμάζουν υποθέσεις ή να διαψεύδουν θεωρίες.

Η απεικόνιση δεδομένων θεωρείται από πολλούς κλάδους ως ένα σύγχρονο ισοδύναμο της οπτικής επικοινωνίας. Περιλαμβάνει τη δημιουργία και τη μελέτη της οπτικής αντιπροσώπευσης των δεδομένων.

Η οπτικοποίηση δεδομένων αναφέρεται στις τεχνικές που χρησιμοποιούνται για την επικοινωνία δεδομένων ή πληροφοριών με την κωδικοποίησή τους ως οπτικά αντικείμενα (για παράδειγμα, σημεία ή γραμμές) που περιέχονται σε γραφικά. Ο στόχος είναι η επικοινωνία των πληροφοριών με σαφήνεια και αποτελεσματικότητα στους χρήστες. Είναι ένα από τα βήματα στην ανάλυση δεδομένων ή την επιστήμη των δεδομένων.

Σύμφωνα με τα παραπάνω η ανάλυση και οπτικοποίηση δεδομένων μπορεί να χρησιμοποιηθεί σε συστήματα επιχειρηματικής ευφυΐας, διότι η Επιχειρηματική ευφυΐα περιλαμβάνει τις στρατηγικές και τις τεχνολογίες που χρησιμοποιούν οι επιχειρήσεις για την ανάλυση δεδομένων των επιχειρηματικών πληροφοριών (είναι ένα από τα τρία βασικά τμήματα της βιομηχανίας πληροφοριών). Οι τεχνολογίες Επιχειρηματικής ευφυΐας προσφέρουν ιστορικές, τρέχουσες και προγνωστικές απόψεις των επιχειρήσεων.

Η ανάλυση και οπτικοποίηση δεδομένων διασυνδέεται με τις λειτουργίες των τεχνολογιών επιχειρηματικής ευφυΐας, δηλαδή τις Επιχειρηματικές αναφορές, την Ηλεκτρονική αναλυτική επεξεργασία, την Εξόρυξη δεδομένων, την Σύνθετη επεξεργασία συμβάντων, την Διαχείριση των επιδόσεων των επιχειρήσεων, την Συγκριτική αξιολόγηση, τις Προγνωστικές αναλύσεις και τις Συντακτικές αναλύσεις.

Άρα η οπτικοποίηση δεδομένων μπορεί να αλλάξει όχι μόνο τον τρόπο με τον οποίο γίνονται αντιληπτά τα δεδομένα αλλά το πόσο γρήγορα και αποτελεσματικά μπορεί να λαμβάνονται αποφάσεις. Ευτυχώς οι λύσεις απεικόνισης εξελίσσονται τόσο γρήγορα όσο και η υπόλοιπη τεχνολογική πρόοδος. Τα διαγράμματα, τα βίντεο, τα πληροφοριακά στοιχεία ακόμη και οι παρουσιάσεις της εικονικής πραγματικότητας και της επαυξημένης πραγματικότητας προσφέρουν όλο και πιο συναρπαστικά και διαισθητικά κανάλια επικοινωνίας.

Βιβλιογραφία

Ampercent - Heatmap. (2018). *A Heatmap Of Sponsored Results on Google SERP's – Second Ad Beats the First?* Ανάκτηση 1 2019, από <https://www.ampercent.com/heatmap-of-sponsored-results-google/9285/>

Blitz, S. (2017). *Exploratory and Confirmatory Analysis: What's the Difference?* Ανάκτηση 12 2018, από <https://www.sisense.com/blog/exploratory-confirmatory-analysis-whats-difference/>

Business Analytics. (2018). *Business Analytics*. Ανάκτηση 1 2019, από <https://practicalanalytics.co/bianalytics-basics/>

CERN - European Laboratory for Particle Physics. (2014). *PAW: Physics Analysis WorkStation*. Ανάκτηση 1 2019, από <http://paw.web.cern.ch/paw/>

CERN-ROOT Data Analysis Framework. (2018). *ROOT*. Ανάκτηση 1 2019, από <https://root.cern.ch/>

Cook, R. (2014, 1 22). *Real Time Reporting For Effective Management*. Ανάκτηση 1 2019, από Toolbox: <https://it.toolbox.com/blogs/erpdesk/real-time-reporting-for-effective-management-012214>

DataMelt Computation & Visualisation. (2018). *DataMelt*. Ανάκτηση 12 2018, από <https://jwork.org/dmelt/>

Datawrapper. (2019). Ανάκτηση 2 2019, από <https://app.datawrapper.de/map/wNWOW/basemap>

DevInfo . (2018). *DevInfo* . Ανάκτηση 12 2018, από <http://devinfo.org/>

DrasticData - treemap. (2018). *DrasticData - treemap*. Ανάκτηση 1 2018, από <https://www.drasticdata.nl/treemapping.htm?d=20151227WDI2014.csv&sc=Air%20transport,%20passengers%20carried&gb=Region&lc=Name&wr=1&dy=1&el=0>

ELKI Data Mining. (2018). *ELKI: Environment for Developing KDD-Applications Supported by Index-Structures*. Ανάκτηση 12 2018, από <https://elki-project.github.io/>

Elmasri, R., & Navathe, S. (2016). *Θεμελιώδεις Αρχές Συστημάτων Βάσεων Δεδομένων* (7η εκδ.). Αθήνα: Διάυλος.

fusioncharts. (2019). Ανάκτηση 2 2019, από [fusioncharts](https://www.fusioncharts.com/): <https://www.fusioncharts.com/>

García deSoto, B., & Adey, B. (2016). Preliminary Resource-based Estimates Combining Artificial Intelligence Approaches and Traditional Techniques. *Procedia Engineering* , 164, σσ. 261-268.

Highcharts. (2019). Ανάκτηση 2 2019, από <https://www.highcharts.com/>

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

Hoelscher, J. M. (2018, 9). Using Tableau to visualize data and drive decision-making. *Journal of Accounting Education* , 44, σσ. 49-59.

Jaan-Matti, S. (2013, 1 3). *19 Things We Can Learn From Numerous Heatmap Tests*. Ανάκτηση 1 2019, από <https://conversionxl.com/blog/19-things-we-can-learn-from-numerous-heatmap-tests/>

KNIME. (2018). *KNIME for Developers*. Ανάκτηση 12 2018, από <https://www.knime.com/>

Król, A. (2016). Transportation Research Procedia. *The Application of the Artificial Intelligence Methods for Planning of the Development of the Transportation Network* , 14, σσ. 4532-4541.

Liang, T.-P., & Liu, Y.-H. (2018, 11). Research Landscape of Business Intelligence and Big Data analytics: A bibliometrics study. *Expert Systems with Applications* , 111, σσ. 2-10.

Microsoft. (2019). Ανάκτηση 2 2019, από Microsoft Power: <https://powerbi.microsoft.com/en-us/>

pandas. (2018). *Python Data Analysis Library*. Ανάκτηση 1 2019, από <https://pandas.pydata.org/>

Past of the Future. (2018). *Past*. Ανάκτηση 12 2018, από <https://folk.uio.no/ohammer/past/>

Plotly. (2019). Ανάκτηση 2 2019, από Modern Analytics Apps for the Enterprise: <https://plot.ly/>

Qlik. (2019). Ανάκτηση 2 2019, από Qlik Analytics Platform: <https://www.qlik.com/us/products/qlik-analytics-platform>

Radenković, M., Lukić, J., Despotović-Zrakić, M., Labus, A., & Bogdanović, Z. (2018, 4). Harnessing business intelligence in smart grids: A case of the electricity market. *Computers in Industry* , 96, σσ. 40-53.

Ramakrishnan, R., & Gehrke, J. (2012). *Συστήματα Διαχείρισης Βάσεων Δεδομένων*, (3η εκδ.). Αθήνα: ΕΚΔΟΣΕΙΣ Α. ΤΖΙΟΛΑ & ΥΙΟΙ Α.Ε.

Russell, S., & Norvig, P. (2005). *Τεχνητή Νοημοσύνη: Μια Σύγχρονη Προσέγγιση* (2η εκδ.). Αθήνα: Εκδόσεις Κλειδάριθμος ΕΠΕ.

SciPy. (2019). *SciPy library*. Ανάκτηση 1 2019, από <https://scipy.org/scipylib/#scipy-library>

Sewell, D. (2018, 12). Visualizing data through curvilinear representations of matrices. *Computational Statistics & Data Analysis* , 128, σσ. 255-270.

Shaulsky, G. (2018). *Orange*. Ανάκτηση 12 2018, από <https://orange.biolab.si/>

Sisense. (2019). Ανάκτηση 2 2019, από <https://www.sisense.com/>

Tableau. (2019). Ανάκτηση 2 2019, από <https://www.tableau.com/>

Εργαλεία ανάλυσης και οπτικοποίησης δεδομένων σε συστήματα επιχειρηματικής ευφυΐας

Tableau. (2019). Ανάκτηση 3 2019, από Tableau Software: <https://public.tableau.com/s/>

The R Foundation. (2018). *The R Project for Statistical Computing*. Ανάκτηση 1 2019, από <https://www.r-project.org/>

Βλαχάβας, Ι., Κεφαλάς, Π., Βασιλειάδης, Ν., Κόκκορας, Φ., & Σακελλαρίου, Η. (2011). *Τεχνητή Νοημοσύνη* (3η εκδ.). Θεσσαλονίκη: Εκδόσεις Πανεπιστημίου Μακεδονίας.

Γιάκας, Α. (2013). *Διεπαφές Μέσω Web για μη Ομοιογενή Σύνολα Δεδομένων σε Πραγματικό Χρόνο*. Ανάκτηση 1 2019, από Πανεπιστήμιο Θεσσαλίας-Πολυτεχνική Σχολή - Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών: <http://ir.lib.uth.gr/handle/11615/42755>

Γναρδέλλης, Χ. (2009). *Ανάλυση δεδομένων με το PASW Statistics 17.0* (2η εκδ.). Αθήνα: ΕΚΔΟΣΕΙΣ ΠΑΠΑΖΗΣΗ ΑΕΒΕ.

Καρλής, Δ. (2005). *Πολυμεταβλητή στατιστική ανάλυση*. Αθήνα: Σταμούλη Α.Ε.

Κύρκος, Ε. (2015). *Επιχειρηματική Ευφυΐα και Εξόρυξη Δεδομένων*. (Σ. Παπαβασιλείου, Επιμ.) Αθήνα: Ελληνικά Ακαδημαϊκά Ηλεκτρονικά Συγγράμματα και Βοηθήματα - Αποθετήριο "Κάλλιπος".

Μπακόλας, Ι., & Τουρνόγλου, Δ. (2013). *Οπτικοποίηση Δεδομένων Ηλεκτρονικής Διακυβέρνησης*. Σάμος: Τμήμα Μηχανικών Πληροφοριακών & Επικοινωνιακών Συστημάτων-Πανεπιστήμιο Αιγαίου.

Ξυνταράκης, Μ., & Αντωνίου, Κ. (2019). Data Science and Data Visualization. *Mobility Patterns, Big Data and Transport Analytics-Tools and Applications for Modeling*, σσ. 107-144.

Παπαδημητρίου, Γ. (2007). *Η ανάλυση δεδομένων*. Αθήνα: Τυπωθήτω.

Παπαδημητρίου, Γ. (2007). *Η Ανάλυση Δεδομένων-Παραγοντική Ανάλυση των Αντιστοιχιών - Ιεραρχική Ταξινόμηση και άλλες Μέθοδοι*. Αθήνα: ΤΥΠΩΘΗΤΩ / ΔΑΡΔΑΝΟΣ.

Ράτσης, Κ. (2009). *Γραφικές Μέθοδοι Παρουσίασης Πολυδιάστατων Δεδομένων*. Πειραιάς: Τμήμα Στατιστικής και Ασφαλιστικής Επιστημής-Πανεπιστήμιο Πειραιώς.

Σιάρδος, Γ. (2017). *Χρηστικό εγχειρίδιο στατιστικής*. Αθήνα: iWrite.gr.

Σταλίδης, Γ., & Καρδαράς, Δ. (2015). *Διαχείριση δεδομένων και επιχειρηματική ευφυΐα*. Αθήνα: Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών.

Χαριδάκου - Κεραμίδα, Δ. (2013). *Εφαρμογή Οπτικοποίησης Αποθηκών Δεδομένων*. Αθήνα: Εθνικό Καποδιστριακό Πανεπιστήμιο Αθηνών-Σχολή Θετικών Επιστημών-Τμήμα Πληροφορικής και Τηλεπικοινωνιών.

Χλουβεράκης, Γ. (2012). *Εισαγωγή στη στατιστική*. Αθήνα: Πεδίο.

Πνευματικά δικαιώματα

Copyright © ΤΕΙ Δυτικής Ελλάδας. Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Δηλώνω ρητά ότι, σύμφωνα με το άρθρο 8 του Ν. 1599/1988 και τα άρθρα 2,4,6 παρ. 3 του Ν. 1256/1982, η παρούσα εργασία αποτελεί αποκλειστικά προϊόν προσωπικής εργασίας και δεν προσβάλλει κάθε μορφής πνευματικά δικαιώματα τρίτων και δεν είναι προϊόν μερικής ή ολικής αντιγραφής, οι πηγές δε που χρησιμοποιήθηκαν περιορίζονται στις βιβλιογραφικές αναφορές και μόνον.

Διένης Παναγιώτης, Μπολοβίνος Αλέξανδρος, [2019]